



# HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı

Eğitimde Ölçme ve Değerlendirme Programı

BİLGİSAYARDA BİREYSELLEŞTİRİLMİŞ TESTLERDE MADDE CEVAPLAMA SÜRESİ,  
YETENEK DÜZEYİ VE MADDE PARAMETRELERİ ARASINDAKİ İLİŞKİNİN İNCELENMESİ

Emine Gülen ULUSOY

Doktora Tezi

Ankara, 2023

Liderlik, arařtırma, inovasyon, kaliteli eđitim ve deđiřim ile

*Daha ileriye ... En İyiyeye ...*



Eğitim Bilimleri Ana Bilim Dalı  
Eğitimde Ölçme ve Değerlendirme Programı

BİLGİSAYARDA BİREYSELLEŞTİRİLMİŞ TESTLERDE MADDE CEVAPLAMA SÜRESİ,  
YETENEK DÜZEYİ VE MADDE PARAMETRELERİ ARASINDAKİ İLİŞKİNİN İNCELENMESİ

INVESTIGATING THE CORRELATIONS BETWEEN RESPONSE TIME, ABILITY LEVEL  
AND ITEM PARAMETERS IN COMPUTER ADAPTIVE TEST

Emine Gülen ULUSOY

Doktora Tezi

Ankara, 2023

## Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,

Emine G¼len ULUSOY'un hazırladıđı "Bilgisayarda Bireyselleřtirilmiř Testlerde Madde Cevaplama S¼resi Yetenek D¼zeyi ve Madde Parametreleri Arasındaki İliřkinin İncelenmesi" bařlıklı bu alıřma j¼rimiz tarafından **Eđitim Bilimleri Ana Bilim Dalı, Eđitimde Ölme ve Deđerlendirme Bilim Dalında Doktora Tezi** olarak kabul edilmiřtir.

J¼ri Bařkanı Prof. Dr. H¼lya KELECİOđLU İmza

J¼ri Üyesi (Danıřman) Prof. Dr. Selahattin GELBAL İmza

J¼ri Üyesi Prof. Dr. Nuri DOđAN İmza

J¼ri Üyesi Do. Dr. Eren Can AYBEK İmza

J¼ri Üyesi Do. Dr. Murat Dođan řAHİN İmza

Enstit¼ Y¼netim Kurulunun  
.../.../.... Tarihli ve .....  
sayılı kararı.

Bu tez Hacettepe niversitesi Lisans¼st¼ Eđitim, đretim ve Sınav Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri yeleri tarafından .... / .... / ..... tarihinde uygun g¼r¼lm¼ř ve Enstit¼ Y¼netim Kurulunca .... / .... / ..... tarihi itibarıyla kabul edilmiřtir.

Prof. Dr. İsmail Hakkı MİRİCİ  
Eđitim Bilimleri Enstit¼s¼ M¼d¼r¼

## Öz

Bu araştırma, Bilgisayarda Bireyselleştirilmiş Testlerde (BBT) madde cevaplama süresinin, yetenek düzeyi ve madde parametreleri ile ilişkisini inceleme amacıyla yapılmıştır. Bu amaç doğrultusunda Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı'nın (ALES) 2017 sonbahar, 2018-1/2/3 dönem soruları ile gerçek zamanlı BBT hazırlanmıştır. Bu dönemlerde sınava girmiş olan adaylardan madde tepki kuramı (MTK) modellerinden Rasch Modeli ile madde ve yetenek parametreleri elde edilmiştir. Bu madde parametreleri ile hazırlanan BBT'nin sayısal testini 238, sözel testini ise 203 birey tamamlamıştır. BBT uygulaması, ALES'e girmeye uygun olan lisans son sınıf öğrencilerine ve lisans mezunu olan bireylere uygulanmıştır. BBT esnasında bireylerin her bir maddeyi cevaplamak için harcadığı süre kaydedilmiştir. Madde cevaplarken harcanan sürenin madde ve yetenek kestirimine dahil edildiği lognormal madde cevaplama süresi modeli (LNCSM) analizleri R Studio'da LNIRT paketi yardımıyla yapılmıştır. LNCSM ile elde edilen madde ve yetenek parametrelerinin MTK'nın Rasch Modeli ve 2 parametrelili lojistik modeli (2PLM) ile elde edilen madde ve yetenek parametreleri ile ilişkisi incelenmiştir. Cevaplama süresini dikkate almayan MTK ile elde edilen madde parametreleri ile LNCSM ile elde edilen madde parametreleri arasında ilişki olma durumları MTK modeline göre (Rasch-2PLM) değişse de süre parametreleri ile madde parametreleri arasındaki anlamlı ilişki bulunmuştur. LNCSM ile kestirilen hız parametresi ile yetenek değerleri arasında negatif ilişki çıkmıştır. Maddeye ve testin toplamına harcanan sürelerin özellikle düşük ve yüksek yetenek grubundaki bireyler için etkili olduğu sonucuna ulaşılmıştır. Ayrıca, LNCSM ile MTK modellerinden elde edilen yetenek ve hız değerleri arasında da ilişki elde edilmiştir.

**Anahtar sözcükler:** bilgisayarda bireyselleştirilmiş test, madde tepki kuramı, cevaplama süresi, lognormal madde cevaplama süresi modeli

## Abstract

This research was conducted to examine the relationship between item response time, ability level and item parameters in Computerized Adaptive Test (CAT). For this purpose, real-time CAT was prepared with the 2017 autumn, 2018-1/2/3 semester questions of the Academic Personnel and Graduate Education Entrance Examination (ALES). Item and ability parameters were obtained from the candidates who took the exam during these periods using the Rasch Model, one of the item response theory (IRT) models. 238 individuals completed the numerical test and 203 individuals completed the verbal test of the CAT prepared with these item parameters. CAT was applied to senior undergraduate students and undergraduate graduates who were eligible to take ALES. During the CAT, the time individuals spent to answer each item was recorded. Lognormal item response time model (LNIRT) analyses, in which the time spent answering items is included in item and ability estimation, were conducted in R Studio with the help of the LNIRT package. The relationship between item and ability parameters obtained with LNIRT and item and ability parameters obtained with IRT's Rasch Model and 2 parameters logistic model (2PLM) was examined. Although the relationship between the item parameters obtained with the IRT, which does not consider the response time, and the item parameters obtained with the LNIRT varies according to the IRT model (Rasch-2PLM), a significant relationship was found between the time parameters and the item parameters. There was a negative relationship between the speed parameter estimated by LNIRT and ability values. It was concluded that the time spent on the item and the total test was effective, especially for individuals in low and high ability groups. Additionally, a relationship was obtained between the ability and speed values obtained from the LNIRT and MTK models.

**Keywords:** computerized adaptive test, item response theory, response time, lognormal response time item response theory model

## Teşekkür

Doktora eğitimim ve tez sürecim boyunca yol göstericiliği, kıymetli fikirleri ve her zaman verdiği destek için; bir “hoca” ve akademisyen nasıl olunur öğrettiği için; son ana kadar bana olan inancı için değerli danışmanım Prof. Dr. Selahattin GELBAL'a çok teşekkür ederim.

Eğitim hayatımda yolumun kesiştiği ve akademik disiplini her zaman örnek almaya çalıştığım Prof. Dr. Hülya KELECİOĞLU'na, sorduğum her soruya sabırla cevap veren kıymetli hocam Prof. Dr. Nuri DOĞAN'a, araştırmamın uygulama sürecinde yardımını esirgemeyen Doç. Dr. Eren Can AYBEK'e, yapıcı yorumları ile bakış açımı geliştiren Doç. Dr. Murat Doğan ŞAHİN'e tez jürimde yaptıkları katkılar nedeniyle teşekkür ederim.

Ders dönemim boyunca yaptığım yolculuklarda kaçırdığım otobüsleri bekleten, yolculuk esnasında çalışmam için gerekli desteği sağlayan otobüs firmalarının muavinlerine teşekkür ederim.

Çalışmamı kolaylaştırmak için her türlü imkanı sağlayan Marmara Üniversitesi'ndeki mesai arkadaşlarıma ve hocalarıma; tezimde karşılaştığım zorlukları problem çözmeci ve eleştirel düşünmeci yaklaşımıyla aşmamı sağlayan Dr. Öğr. Üyesi Orhan ÇANAKÇI'ya; tez sürecimdeki bütün sancularıma ve sevinçlerime ortak olan Dr. Zehra Betül ALP'e sonsuz teşekkürlerimi sunarım.

İhtiyacım olduğunda yardımını ve kıymetli fikirlerini esirgemeyen Dr. Eren Halil ÖZBERK'e ve Doç. Dr. Sümeyra SOYSAL'a çok teşekkür ederim. Çalışabilmem için her türlü imkanı sağlayan ve bana rehberlik eden Semra GELBAL'a desteği için çok teşekkür ederim. Doktora eğitimimi tamamlamam için motivasyonumu kaybetmememi sağlayan ve desteklerini esirgemeyen bütün arkadaşlarıma teşekkür ederim, iyi ki varsınız!

Tezimi yazarken bunaldığım zamanlarda beni güldürmeyi başaran yeğenim Mustafa Şafak'a; fikirlerine her zaman güvendiğim ve danıştığım Kemal ÜNLÜ'ye; hayatta her şeye ortak olan kardeşim Ayşenur'a; beni her zaman cesaretlendiren anneme; attığım her adımda arkamda olan ve desteğini hep hissettiğim babama (ki tezi senin inancın sayesinde bitirdim);

eđitim hayatımın en bařından beri her yere benimle gelen Ayřeannem'e ok teřekkür ederim. Öğretmenlik genlerimi aldıđım rahmetli Muallim Mehmet Dedem'e, akademik becerilerimin birçođunu ondan aldıđıma inandıđım rahmetli Babaannem'e ve bugünlerimi göremese de beni her zaman izlediđine inandıđım, üzerimde ok emeđi olan rahmetli Mustafa Dedem'e sonsuz teřekkür ederim.



*Beni büyütenlere;  
Şafak'tan Ayşeannem'e...*

## İçindekiler

Kabul ve Onay .....	ii
Öz.....	iii
Abstract .....	iv
Teşekkür.....	v
Tablolar Dizini.....	xi
Şekiller Dizini .....	xii
Simgeler ve Kısaltmalar Dizini .....	xiii
Bölüm 1 Giriş .....	15
Problem Durumu .....	30
Araştırmanın Amacı ve Önemi .....	31
Araştırma Problemi.....	31
Alt Problemler.....	32
Sayıtlılar .....	32
Sınırlılıklar .....	32
Tanımlar .....	33
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar .....	34
Bilgisayarda Bireyselleştirilmiş Testler İle İlgili Yapılmış Çalışmalar.....	34
Testte Cevaplama Süresi İle İlgili Araştırmalara Genel Bir Bakış .....	40
Bölüm 3 Yöntem .....	45
Araştırmanın Türü .....	45
Çalışma Grubu .....	45
Veri Toplama Süreci.....	46
Veri Toplama Araçları.....	46
Verilerin Analizi.....	51
Bölüm 4 Bulgular, Yorumlar ve Tartışma .....	58
Araştırmada Elde Edilen Değerlere İlişkin Betimleyici İstatistikler .....	58

1. Alt Problem “Yetenek düzeyleri ile bireylerin cevapladıkları madde sayıları, madde parametreleri, madde cevaplama süreleri ve test toplam süreleri arasında istatistiksel olarak anlamlı ilişki var mıdır?” .....	61
2. Alt Problem “Düşük ve yüksek yetenek düzeylerine göre bireylerin cevapladıkları madde sayıları, madde parametreleri, madde cevaplama süreleri ve test toplam süreleri arasında istatistiksel olarak anlamlı fark var mıdır?” .....	63
3. Alt Problem “Maddenin doğru veya yanlış cevaplanmasına göre madde cevaplama süresinde istatistiksel olarak anlamlı fark var mıdır?” .....	66
4. Alt Problem “Madde cevaplama süresini dikkate alan lognormal cevaplama süresi modeli (LNCSM) ile kestirilen madde parametreleri ve madde cevaplama süresini dikkate almayan MTK modelleri ile kestirilen madde parametreleri arasında istatistiksel olarak anlamlı ilişki var mıdır?” .....	72
5. Alt Problem “Madde cevaplama süresini dikkate alan lognormal cevaplama süresi modeli (LNCSM) ile kestirilen yetenek düzeyleri ve madde cevaplama süresini dikkate almayan MTK modelleri ile kestirilen yetenek düzeyleri arasında istatistiksel olarak anlamlı ilişki var mıdır?” .....	75
Bölüm 5 Sonuç ve Öneriler .....	78
Sonuçlar .....	78
Öneriler .....	79
Kaynaklar .....	80
EKLER .....	90
Ek-A: Sayısal Testinde Madde Cevaplama Süresinin Maddenin Doğru veya Yanlış Cevaplanma Durumuna Göre Değişim Grafikleri .....	90
Ek-B: Sözel Testinde Madde Cevaplama Süresinin Maddenin Doğru veya Yanlış Cevaplanma Durumuna Göre Değişim Grafikleri .....	94
Ek-C: LNCSM Sonucu Elde Edilen Madde Parametrelerinin Betimsel Analiz Sonuçları .....	98
Ek-D: LNCSM Sonucu Elde Edilen Yetenek ve Hız Parametrelerinin Betimsel Analiz Sonuçları .....	99
Ek-E: Araştırma Etik Komisyonu Onay Bildirimi .....	100

EK-F: Etik Beyanı .....	101
EK-G: Doktora Tez Çalışması Orijinallik Raporu .....	102
EK-H: Dissertation Originality Report .....	103
EK-I: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı .....	104

## Tablolar Dizini

<b>Tablo 1</b> <i>ALES'in Sayısal Maddelerinin Becerilere Göre Dağılımı</i> .....	47
<b>Tablo 2</b> <i>ALES'in Sözel Maddelerinin Becerilere Göre Dağılımı</i> .....	48
<b>Tablo 3</b> <i>Sayısal Testinde Yanıtlanan Madde Sayılarının Frekans Dağılımları</i> .....	48
<b>Tablo 4</b> <i>Sözel Testinde Yanıtlanan Madde Sayılarının Frekans Dağılımları</i> .....	49
<b>Tablo 5</b> <i>Doğrulayıcı Faktör Analizine Ait Model Uyum İndeksleri</i> .....	52
<b>Tablo 6</b> <i>Rasch Model ve 2 PLM İçin Model Veri Uyumu İncelemesi</i> .....	53
<b>Tablo 7</b> <i>Madde Parametrelerinin Değişmezliğine Ait Korelasyonlar</i> .....	55
<b>Tablo 8</b> <i>Yetenek Parametrelerinin Değişmezliğine Ait Korelasyonlar</i> .....	56
<b>Tablo 9</b> <i>Sayısal Testinde Bireylerin Cevapladıkları Sorulara Ait Betimsel Analiz Sonuçları</i> .....	58
<b>Tablo 10</b> <i>Sözel Testinde Bireylerin Cevapladıkları Sorulara Ait Betimsel Analiz Sonuçları</i> .....	59
<b>Tablo 11</b> <i>Sayısal Testinde Maddelere Ait Betimsel Analiz Sonuçları</i> .....	60
<b>Tablo 12</b> <i>Sözel Testinde Maddelere Ait Betimsel Analiz Sonuçları</i> .....	60
<b>Tablo 13</b> .....	61
<i>Sayısal Testinde Bireylerden Elde Edilen Değişkenler Arasındaki İlişkiler</i> .....	61
<b>Tablo 14</b> <i>Sözel Testinde Bireylerden Elde Edilen Değişkenler Arasındaki İlişkiler</i> ...	62
<b>Tablo 15</b> <i>Sayısal Testinde Yetenek Düzeyine Göre Değişkenlere Ait Bağımsız Örneklem t Testi Tablosu</i> .....	64
<b>Tablo 16</b> <i>Sözel Testinde Yetenek Düzeyine Göre Değişkenlere Ait Bağımsız Örneklem t Testi Tablosu</i> .....	65
<b>Tablo 17</b> <i>Sayısal Testinde LNCSM ve MTK Modellerinden Elde Edilen Madde Parametreleri Arasındaki İlişkiler</i> .....	73
<b>Tablo 18</b> <i>Sözel Testinde LNCSM ve MTK Modellerinden Elde Edilen Madde Parametreleri Arasındaki İlişkiler</i> .....	74
<b>Tablo 19</b> <i>Sayısal Testinde LNCSM ve BBT'den Elde Edilen Yetenek ve Hız Değerleri Arasındaki İlişkiler</i> .....	76
<b>Tablo 20</b> <i>Sözel Testinde LNCSM ve BBT'den Elde Edilen Yetenek ve Hız Değerleri Arasındaki İlişkiler</i> .....	77

**Şekiller Dizini**

<b>Şekil 1</b> <i>BBT Algoritmasının Akış Şeması</i> .....	18
<b>Şekil 2</b> <i>İyi model veri uyumu Madde Karakteristik Eğrisi</i> .....	21
<b>Şekil 3</b> <i>Uyumsuz model veri uyumu Madde Karakteristik Eğrisi</i> .....	22
<b>Şekil 4</b> <i>Sayısal Testinde Madde Cevaplama Süresinin Maddenin Doğru veya Yanlış Cevaplanma Durumuna Göre Değişim Grafikleri</i> .....	68
<b>Şekil 5</b> <i>Sözel Testinde Madde Cevaplama Süresinin Maddenin Doğru veya Yanlış Cevaplanma Durumuna Göre Değişim Grafikleri</i> .....	70

## Simgeler ve Kısaltmalar Dizini

BBT: Bilgisayarda Bireyselleştirilmiş Test

MTK: Madde Tepki Kuramı

ÇB-MTK: Çok Boyutlu Madde Tepki Kuramı

ALES: Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı

ÖSS: Öğrenci Seçme Sınavı

a parametresi: Madde ayırt ediciliği

b parametresi: Madde güçlüğü

c parametresi: Şans parametresi

1 PLM: 1 Parametrelili Lojistik Model

2 PLM: 2 Parametrelili Lojistik Model

3 PLM: 3 Parametrelili Lojistik Model

ss: Standart sapma

SH: Standart Hata

$\theta$ : yetenek parametresi

## Bölüm 1

### Giriş

Ölçme; gözlenen, ölçülen nesne ya da özellikleri bir ölçü birimine göre anlatım işlemidir. Buna nesnelere ya da bireylere, belirli özelliğe sahip oluş derecesini belirtmek için, belirli kurallara uyarak simgesel değerler verme işlemi de denir. Değerlendirme; çeşitli ölçme ya da ölçme yerine geçen gözlemler sonucunda yapılan yargıyı belirten genel terimdir ve eğitimde özellikle öğretimde amaçlara ne derece yaklaşıldığının bir ölçüsünü verir (Tekin, 2000). Eğitim alanında belirlenen hedeflere ulaşıp ulaşılmadığını ve ulaşıldı ise hangi oranda ulaşıldığını belirlemek için ölçme değerli bir araçtır.

Eğitimde ölçme yapmak için farklı test türlerinden yararlanılmaktadır. Bir testte yer alan maddelerin tümünün her bireye uygulandığı testler “sabit ve değişmez” olarak tanımlanırken; madde sayılarının, maddelerin bireyler için değişkenlik gösterdiği testler ise “Bireye Uyarlanmış Test” olarak adlandırılır (Raiche ve Blais, 2002). Bu iki test türü arasında testin sunumu, güçlüğü, uzunluğu, uygulama zamanı, süresi ve sonuçlandırması açısından farklılıklar vardır. Geleneksel test uygulamalarında ortalama beceri düzeyindeki bireyler benzer bir testi aynı anda alırlar. Bu testler genellikle uzun ve fazla süre gerektiren testler olup sonuçları anında verilmez. Oysa bireye uyarlanmış test uygulamalarında farklı beceri düzeyindeki bireylerin her biri farklı bir testi kendileri için uygun zamanda alırlar. Bu testler genellikle geleneksel testlere göre kısa ve az süre gerektiren testler olup sonuçları anında verilir (Tian, Miao, Zhu ve Gong, 2007).

İdeal bir test uygulaması, her bir test katılımcısına kendi yetenek düzeyine uygun maddeleri sunmayı hedefler. Ancak, bireyin yetenek düzeyi önceden biliniyorsa, uygun testin uygulanmasının anlamı ortadan kalkar. Bu durumu aşmak için, test uygulamasında her katılımcının yetenek düzeyine bağlı olarak sıralanan uygun maddeleri sırasıyla almaları sağlanır. Bunu yapmak için katılımcının test boyunca verdiği cevaplar kullanılarak bir sonraki madde belirlenir. Bu yöntem, her katılımcının kendi yetenek düzeyine en uygun maddelerle karşılaşmasını sağlayarak test sonuçlarının daha doğru ve güvenilir olmasına yardımcı olur



(Kalender, 2004a'da Weiss'den alıntılındığı gibi, 1983). Bu öneri doğrultusunda eğitimde ölçme alanında geliştirilen bireye uyarlanmış testler, bilgisayarın test uygulamalarında kullanılması ile birlikte "Bilgisayarda Bireyselleştirilmiş Test (BBT)" (Weiss, 2004) olarak adlandırılmıştır. BBT'nin üç önemli ana bileşeni vardır: bilgisayar ortamında olması, bireyselleştirilmiş olması ve test olması. Bilgisayar ortamında yapılan tüm testlerin, bireyselleştirilmiş olmadığını da belirtmek gerekir. Örneğin, bir bilgisayar tabanlı bir testte tüm adaylara aynı maddeler uygulanabilir ve maddelerin zorluğu adaylara göre ayarlanmış veya bireyselleştirilmiş olmayabilir.

BBT'lerin kağıt-kalem testlerine kıyasla pek çok avantajları bulunmaktadır. Bu avantajlar:

- Geleneksel testlere oranla daha az maddeyle uygulanması,
- Yanıtlayıcının farklı teste geçmek için diğer yanıtlayıcıları beklemek zorunda kalmaması,
- Her bir bireyin farklı bir test alması ve testin dijital ortamda saklanması sebebiyle esnek, düzenli ve güvenilir olması,
- Yanıtlayıcı yeteneği ve madde güçlüklerinin eşleştirilebilmesine olanak sağlaması sebebiyle daha geniş aralıkta yetenek düzeyindeki kişilere ulaşılabilmesi,
- Cevaplama süresi, maddede kalma süresi, vb. özelliklere yönelik bilgiler verebilmesi; test sonuçlarının hemen alınması,
- Uygulayıcı kaynaklı ölçüm hatalarından arınık olması

olarak sıralanabilir (Hambleton ve ark., 1991; Kingsbury ve Hauser, 2004; Thompson, 2010; Wainer, 2000).

BOBUT uygulamalarının avantajlarının yanı sıra dezavantajları da bulunmaktadır. Bu dezavantajlar; "nispeten maliyetli uygulamalar olmaları, geniş ölçekli uygulamaların yapılması

istenildiğinde her bölgede aynı teknolojik alt yapının sağlanamaması, donanımsal sınırlılıklar, teknolojik okuryazarlık düzeylerinin sonuçları etkileyebilir olması, her konu ya da yetenek için uygulanmasının zor olması, madde havuzunda çok fazla madde olması gerektiğinden bunu sağlamanın zor olması, geleneksel testlerden elde edilen madde parametrelerinin (örn. madde karakteristik eğrisi) BOBUT ile elde edilenden farklılık göstermesi, her yanıtlayıcıya farklı maddeler gelmesi sebebiyle kişiler üzerinde olumsuz algılar oluşması olarak sıralanabilir (ACRE, 2010; Çıkrıkçı Demirtaşlı, 1999; Embretson ve Reise, 2000; Şenel, 2017).” Fakat tüm bu dezavantajların yanında günümüz dünyasında bilgisayar ve yazılım dünyasında yaşanan gelişmelerle bireye özgü testler bilgisayar ortamında daha etkili şekilde uygulanabilir hale gelmiştir (Özbaşı, 2014).

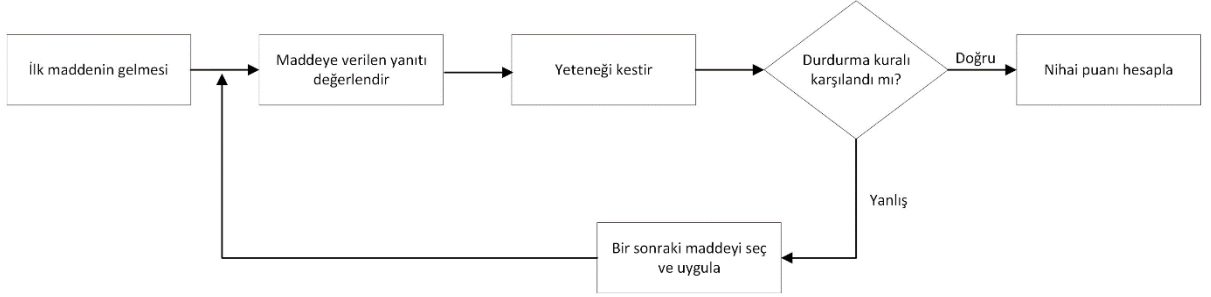
BBT, psikometrik özellikleri bakımından daha önce kalibre edilmiş maddelerden oluşan bir madde havuzundan cevaplayıcıların yetenek ve ilgi düzeylerine uygun maddeler seçerek onlara özgü testler oluşturmaktadır (Weiss, 2004). Bu açıdan BBT, maddeye cevap veren cevaplayıcının önceki cevaplarının dikkate alındığı, karşısına gelecek bir sonraki maddenin madde havuzundan seçildiği bir döngü şeklinde ilerlemektedir. Bireyin maddelere verdiği her cevap sonunda ölçülen özellikler yeniden kestirilir ve bu süreç test sonlanana kadar devam eder (Alkan, 2021). MTK’da madde parametreleri yanıtlayıcılardan, yanıtlayıcıların özellikleri de madde örneklemeden bağımsız şekilde olup her bir yanıtlayıcı için ayrı standart hata kestirimi yapılabilmektedir. Bu nedenle MTK, BBT uygulamaları için oldukça kullanışlı bir modeldir (Yaşar ve Aybek, 2019). BBT, MTK’nın değişmezlik özelliğini kullanarak bir algoritma oluşturur ve bu algoritma ile her yanıtlayıcının kendisi için iyi ölçme yapabilen bir test almasına olanak sağlar (Embretson ve Reise, 2013).

BBT’nin çoğu uygulamasında adaylar testte, her maddeye sırayla cevap vermek durumundadır. Çünkü adaya sunulacak her yeni maddenin seçimi, daha önce cevapladığı maddelerde sunduğu performansa bağlıdır. Bu yüzden; aday bir maddeyi daha sonra cevaplandırmak üzere atlayamaz. Bu, zor bir maddeyi cevaplandırmayı sınavın sonunda tekrar ele almak gibi bazı test tekniklerinin kullanımını sınırlandırır. Aynı şekilde, adaylar cevap

verdikleri maddeyi tekrar değerlendirerek cevaplarını değiştirme hakkına da sahip değildirler. Çisar ve ark. (2010) sunduğu BBT akış şeması şekil 1’de verilmiştir.

### Şekil 1

#### BBT Algoritmasının Akış Şeması



BBT'nin beş özelliği bulunmaktadır (Istiyono ve ark., 2020; Baker, 1983; Weiss, 2011).

Bu özellikler:

- BBT, istatistiksel özellikleriyle tamamlanmış bir dizi test maddesi üzerine kuruludur.
- Test, sınava giren kişinin madde havuzundan bir başlangıç noktası seçmesini kolaylaştırır, böylece tüm sınava giren kişiler aynı maddeyle başlamaz.
- Test puanları, farklı test katılımcılarına verilen farklı test maddelerinden elde edilebilir.
- Sonraki maddelerin seçimi, sınava giren kişinin önceki maddeye verdiği cevaplara bağlıdır.
- Bir test, her birey tarafından yapılan farklı sayıda maddeyle bile belirli sonlandırma ölçütlerine ulaşıldığında sona erer.

#### **Madde Tepki Kuramı (MTK)**

Madde Tepki Kuramı (MTK) modelleri “bir ölçme aracı tarafından ölçülen yetenek ( $\theta$  sembolü ile gösterilen) ya da özellik ile bir maddeye verilen cevap arasındaki ilişkiyi gösterir. Maddeye verilen cevap iki (örn: doğru-yanlış, evet-hayır vb.) ya da çoklu kategorili (iki

kategoriden daha fazla-polytomous) olabilir; örneğin bir hakemin ya da puanlayıcının derecelendirmesi veya Likert tipi derecelendirme ölçekleri. MTK'nın kullanılmasının en temel nedenlerinden biri testleri veya anketleri puanlamaktır. MTK puanları genellikle "yetenek, özellik veya yeterlik" olarak adlandırılır ve hesaplamasında madde güçlüğü ve ayırt ediciliği de dikkate alınır. Bu nedenle MTK puanlarının ham puanlardan daha güvenilir olduğu kabul edilir (DeMars, 2016).

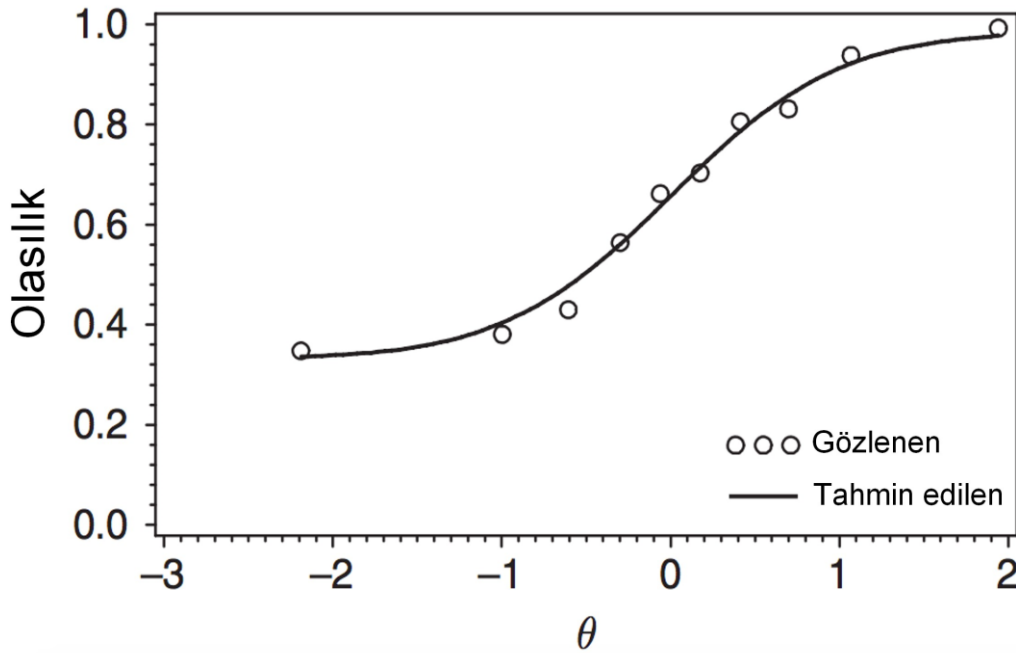
MTK modelleri, fonksiyonların parametrelerine bağlı olarak değişebilir ve aynı zamanda fonksiyon türüne göre de farklılık gösterebilir (Demirel, 2021). MTK'da bulunan "a parametresi – madde ayırt ediciliği", "b parametresi – madde güçlüğü" ve "c parametresi – şans başarısının etkisi" olarak adlandırılır. 1PLM (1 parametrelili lojistik model), MTK modellerinin en temel şekli olarak kabul edilir. Bu model, bir maddeye verilen doğru cevap olasılığını, cevaplayıcının yeteneği ve maddenin zorluğu göz önüne alındığında tahmin eden madde yanıt fonksiyonunu temsil eder. Ayrıca, bu modelde şans parametresi 0 olarak kabul edilirken ayırt edicilik parametresi 1 olarak hesaba katılır (Çörtük, 2022). 1PLM'de, ayırt edicilik parametresi tüm maddeler için sabit bir değere sahiptir ve yetenek ölçeği boyunca farklı maddeler için elde edilen madde karakteristik eğrileri paraleldir. Bir diğer özel MTK modeli olan Rasch Modeli olarak bilinen model, bir cevaplayıcının bir test maddesine doğru cevap verme olasılığını hem cevaplayıcının yeteneği hem de maddenin zorluğu ile ilişkilendiren bir fonksiyon olarak tanımlar (Susmann, 2018). Rasch modeli genellikle yalnızca madde ve birey parametrelerinin tahmin edildiği ve tahmin edilen parametre değerlerinin arasındaki farklılıkları açıklamak için herhangi bir tahmin edici değişkenin kullanılmadığı bir tanımlayıcı bir model olarak kabul edilir (Atar, 2011). Rasch modeli, her madde için yalnızca bir konum parametresi tahmin eder ve ayırım parametresinin tüm maddeler için sabit olduğunu varsayar. Rasch modelinin ana avantajlarından biri, daha küçük örneklemelere uyum sağlayabilmesidir. Ayrıca Rasch modeli, modelin verilere uygun olması durumunda herhangi iki maddenin konum parametrelerinin, ilgili katılımcı grubundan bağımsız olarak karşılaştırılabildiği ve herhangi iki katılımcının yetenek puanının, kullanılan madde setinden bağımsız olarak karşılaştırılabildiği bir modeldir (Nguyen

ve ark., 2014). 1PL ve Rasch modellerini matematiksel olarak eşit olduğu bilinmektedir. Fakat Rasch modeli, madde ayırt ediciliğini ( $a_i$ ) 1 ile sınırlarken, 1PLM, ayırt edicilik parametresini 1 ile sınırlamaz (DeMars, 2010). 2PLM (2 parametrelili lojistik model), iki parametre (güçlük  $b_i$  ve ayırt edicilik  $a_i$ ) kullanarak başarılı bir cevap olasılığını tahmin etmede kullanılır ve şans parametresi, 1PLM'de olduğu gibi, bu modelde de 0 olarak kabul edilir. 3PLM (üç parametrelili lojistik model), Birnbaum (1968)'un 2PLM'ye doğru cevap verme olasılığına şans başarısının etkisini gösteren üçüncü bir parametre ( $c_i$ ) eklemesiyle geliştirilmiştir. Bu model, 2 PLM'nin özel bir hali olarak kabul edilmektedir (Selçuk, 2023).  $c_i$  şans parametresi, düşük yetenekli bir bireyin  $i$  maddesini şansla doğru cevaplama olasılığını göstermektedir. Madde karakteristik eğrisi ile x eksenini arasındaki uzaklık şans parametresine işaret etmektedir (Embretson ve Reise, 2000).

MTK temelde, "tek boyutluluk, yerel bağımsızlık ve uygun model özellikleri" olmak üzere 3 varsayıma dayanmaktadır. Tek boyutluluk; "tek bir örtük özelliğin ölçülmesi" olarak tanımlanabilir. Bazı durumlarda bu varsayımın karşılanması mümkün olmayabilir. Bu nedenle bir örtük özelliği aramak yerine baskın bir özellik aranır (Hambleton ve ark., 1991). Bu varsayımın karşılanma durumu faktör analizi sonucunda elde edilen maddeler arası korelasyon matrisinin öz değerleri incelenerek kontrol edilebileceği gibi (DeMars, 2010), Stout'un temel tek boyutluluk testi ve tek boyutluluk çözümündeki artıklara dayanan indeksler incelenerek de kontrol edilebilir (DeMars, 2016). İkinci varsayım olan yerel bağımsızlık ise; "test performansı üzerinde etkili olan yetenek düzeyi sabit tutulduğunda, bireylerin maddelere vereceği yanıtların birbirinden bağımsız veya ilişkisiz olma" durumudur (Selçuk, 2023). Yerel bağımsızlık varsayımının karşılanması için ilk varsayım olan tek boyutluluk varsayımının karşılanmış olması gerekir (Hambleton ve ark., 1991). Yen (1984) madde çiftlerinin yerel bağımsızlığını kontrol etmek için  $Q_3$  testini önermektedir.

## Şekil 2

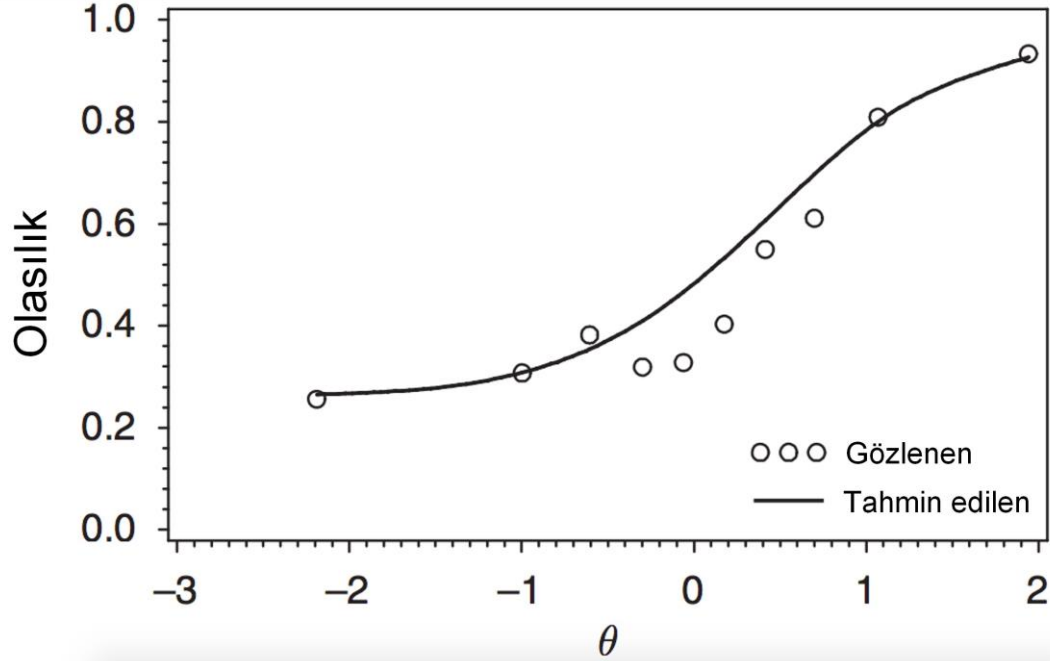
*İyi model veri uyumu Madde Karakteristik Eğrisi*



DeMars (2016) üçüncü varsayım olan “uygun model özellikleri”nde ise modelin yanlış tanımlanmasını kontrol etmek için model ve veri arasındaki uyumun belirlenebileceğine değinmektedir. Örneğin; “veriler farklı eğimlerde ya da sıfırdan farklı düşük asimptotu olan bir modeli gösterirse ve 1PLM kullanılırsa, maddelerin çoğu 1PLM ile uyumlu olmaz. Eğer fonksiyon monoton olarak artmıyorsa yaygın olarak kullanılan modellerin hiçbiri uyum sağlamaz (DeMars, 2016).” Madde uyumunda, “artık ve gözlenen kavramı” ve “modelden kestirilen (beklenen) oran kavramı” kilit kelimelerdir. Artıklar  $\theta$ 'ya koşullu olmakla birlikte bu artıklar yaklaşık aynı  $\theta$ 'ya sahip bireyler için hesaplandığını açıklar. Şekil 2'de iyi bir model veri uyumuna sahip madde karakteristik eğrisi görülürken, şekil 3'te ise model veri uyumu düşük olan bir madde karakteristik eğrişi görülmektedir (DeMars, 2016, s.51). Eğer model uyumluysa her bir  $\theta$  düzeyinde gözlenen doğru cevap oranının model beklentisine yakın olması beklenir.

### Şekil 3

Uyumsuz model veri uyumu Madde Karakteristik Eğrisi



MTK, model tabanlı ölçümler için metodolojik bir temel sağlamaktadır. MTK'da madde yanıtları, bireysel özellik düzeylerinin ve madde özelliklerinin bir fonksiyonu olarak modellenir. Madde güçlükleri ve bireysel özellik düzeyleri ortak bir ölçekte açıklanabilir (Hartig ve Höhler, 2009). Fakat günümüz psikolojik yapılar düşünüldüğünde bu yapıların çok boyutlu olduğu görülmektedir (Şahin, 2017). Bu açıdan “Çok Boyutlu Madde Tepki Kuramı (Multidimensional item response theory – MIRT)” psikolojik ölçüm için ana akım bir temel oluşturmaktadır (Embretson ve Reise, 2000). Çok Boyutlu MTK (ÇB-MTK), karmaşık alanlarda performansı yeterince modelleme potansiyeline sahip olmakla birlikte bireysel test maddeleri için gerekli olan yetenek karışımları ile bile birden fazla yetenek ÇB-MTK'da aynı anda dikkate alınabilir. Bu sebeple ÇB-MTK eğitim alanındaki yeterlilikleri değerlendirmede oldukça kullanışlı bir yöntemdir (Hartig ve Höhler, 2009). ÇB-MTK'lar “telafi edici (compensatory)” ve “telafisel olmayan ya da kısmi telafisel (non-compensatory – partially compensatory)” olmak üzere iki grupta incelenmektedir (Şahin, 2017). Telafi edici ÇB-MTK'da bir boyuttaki düşük bir yetenek, ikinci bir boyuttaki yüksek bir yetenekle telafi edilebilir. Telafisel olmayan bir modelde ise, belirli

bir madde için gereken tüm yetenekler yüksekse, başarı olasılığı yalnızca bire yaklaşacaktır. Daha basit bir ifade ile kişinin bir boyuttaki yeterliği diğer bir boyuttaki yetersizliğini telafi edemez (Hartig ve Höhler, 2009). ÇB-MTK modellerinde ölçümün tipik amacı, belirli bir içerik alanında başarılı bir şekilde performans göstermek için gerekli olan çok sayıda farklı yeteneği değerlendirmektir. Bu sebeple ÇB-MTK modellerinde klasik MTK modellerinin varsayımlarından biri olan “*tek boyutluluk*” varsayımı aranmamaktadır.

### ***Bilgisayarda Bireyselleştirilmiş Test***

Son yıllarda, özellikle geniş ölçekli testlerin geliştirilmesi ve değerlendirilmesinde madde tepki kuramı (MTK) modellerinden yararlanılmaktadır (C.AT. Central, 2010). Bu kuramın modelleri klasik test kuramına göre farklı ölçek nitelikleri vermekte, bu da elde edilen test puanlarının daha karşılaştırılabilir nitelikte olmasını sağlamaktadır (Yıldırım, Çömlekoğlu ve Berberoğlu, 2003). BBT’yi mümkün kılan iki ana yenilikten biri bilgisayar teknolojisi diğeri de MTK’dır. BBT’nin kağıt-kalem testlerine karşı tercih edilmesini sağlayan iki kritik özelliği vardır. Bunlardan ilki, MTK’da tüm test yerine maddelerin tek tek analiz edilmesi ve böylece BTT uygulamasında adayın yeteneğinin kestirimi için her maddenin ayrı bir fonksiyon sergileyebilmesidir. Maddelerden bağımsız olarak, maddeleri cevaplayanları aynı sürekli ölçeğe yerleştirdiği için testin, adayların yeteneğine göre uyarlanması da diğer bir kritik özelliktir (Smith, 2000). Buna göre, adaylar çoklu zorluk seviyelerinde farklı maddeler içeren testler alsalar dahi; performansları, MTK ölçeklendirme uygulamalarıyla tahmin edilebilir ve kıyaslanabilir.

BBT, “MTK’nın değişmezlik özelliğini kullanarak oluşturduğu algoritma ile her bir cevap verenin kendisi için iyi bir ölçme yapabilen bir test almasına olanak sağlayan bir MTK uygulamasıdır (Embretson ve Reise, 2000).” BBT uygulamaları, teste katılan bireylerin her birine kendi yetenek ve ilgilerine bağlı olarak maddeleri yanıtlama imkanı sunmaktadır (Alkan, 2021). Kişinin yetenek ve ilgi düzeyine göre maddelerin gelmesi sebebiyle başarı düzeyi yüksek olan bireyler daha zor maddeleri cevaplarken, başarı düzeyi düşük olanlar ise kendi yetenek düzeylerine uygun şekilde daha kolay maddeler ile karşılaşmaktadırlar. Bu bakımdan



BOBUT uygulama sonunda her bir bireye özel form oluşturmaktadır (Bartram ve Hambleton, 2006; De Ayala, 2009; DeMars, 2010).

BBT uygulaması sırasında adaya performansına uygun olarak maddeler yöneltilir. Teste katılan bireylerin her birine kendi yetenek ve ilgilerine bağlı olarak maddeleri yanıtlama imkanı sunan BBT uygulamalarında (Alkan, 2021) başarı düzeyi yüksek olan bireyler daha zor maddeler ile karşılaşırken başarı düzeyi düşük olanlar ise kendi yetenek düzeylerine uygun şekilde daha basit maddelerle karşılaşmaktadırlar. Bu nedenle BBT uygulama sonunda her bir bireye özel form oluşturmaktadır (Bartram ve Hambleton, 2006; De Ayala, 2009; DeMars, 2010). BBT uygulamasında her bir kişi farklı maddeleri yanıtlasa bile yanıtlayıcıların yetenek kestirimlerinin karşılaştırılması mümkün olmaktadır (Şenel, 2017). MTK ölçeklendirmesi kullanımı, düşük ve yüksek yetenek seviyesindeki adayların aynı ölçeğe yerleştirilmesini sağladığı için adaylar aynı maddeleri almamış olsalar dahi, yetenek tahminleri karşılaştırılabilir.

BBT uygulamalarında, sonuçları etkileyen birçok farklı durum vardır. Bunlar testi başlatma kuralı, madde seçme yöntemi, yetenek kestirim yöntemi, içerik dengeleme yöntemi ve testi sonlandırma kuralı olarak sıralanabilir.

Yeterli madde havuzu oluşturduktan sonra karar verilecek ilk şey BBT'nin nasıl başlayacağıdır. Bireylerin yetenek düzeyleri hakkında bilgi sahibi olunmadığı zaman sıklıkla kullanılan yöntem orta düzey zorlukta madde ile başlanmasıdır (Embretson ve Reise, 2000).

Madde seçme yöntemi, birey testi alırken sıradaki maddenin nasıl seçileceğinin belirlenmesidir. Bu yöntemlerden bazıları Maksimum Fisher Bilgi, Kullback-Leibler Uzaklığı ve Beklenen Maksimum Bilgidir.

Maksimum Fisher Bilgi yöntemi, birden fazla parametre için Fisher bilgi matrisini dikkate alan ve onu çok boyutlu değerlendirmeler için uygun hale getiren, madde havuzundan maddelerin yinelemeli olarak seçilmesini içerir. Seçim süreci, istenen test uzunluğu veya hedef

ölçüm hassasiyeti seviyesi gibi önceden belirlenmiş bir kriter karşılanıncaya kadar devam eder (Van der Linden ve Glas, 2010).

Kullback-Leibler Uzaklığı madde seçim yönteminde, mevcut ve istenen madde zorluk dağılımı arasındaki farklılığı ölçmek için Kullback-Leibler Uzaklığı kullanır. Bu tutarsızlığın azaltılmasına en fazla katkıda bulunan maddeler uygulama için seçilir (Van der Linden ve Glas, 2010).

Beklenen Maksimum Bilgi yöntemi bireyin tahmini yetenek düzeyinde madde karakteristik eğrisinin en dik eğimine sahip maddeleri seçmeyi amaçlar. Bu yöntem, testin, testi yapan kişinin yeteneğine hızlı bir şekilde uyum sağlamasını sağlayarak etkili ve hassas ölçüm yapılmasını sağlar (Baker, 2001).

Yetenek kestiriminde kullanılan başlıca iki yöntem Maksimum Olabilirlik Yöntemi ve Bayes Yöntemidir. Maksimum olabilirlik yönteminde, her bir madde düzeyindeki mevcut tüm bilgi göz önüne alınır. Bu nedenle, birey zor bir maddeyi doğru cevapladığında, yetenek kestirimindeki değişim, kolay bir maddeyi doğru cevapladığında gözlenen değişimden daha büyük olacaktır (Weiss, 2011). Bayes Yönteminde yetenek parametrelerine ait önsel bir dağılım kullanılır. Yetenek parametrelerine ait önsel dağılımın bilindiği durumlarda ise sonsal dağılım hesaplanır. Yetenek parametresine ait sonsal dağılım, bir sonraki maddeye verilecek yanıtın önsel dağılımı olarak ele alınır. Bireye ait son yetenek kestirimi gerçekleşinceye kadar bu döngü devam eder (Song, 2010).

BBT yer alan maddelerin hepsinin benzer sıklıkta bireylere gelmesi için içerik dengeleme yöntemlerine başvurulur. A-Stratified – Tabakalı dengeleme yöntemi düşük a değerli maddelerin kullanımını artırarak madde kullanım sıklığı dağılımını eşitlemeye çalışan a-tabakalı tasarımıdır. Yöntemin belirli madde havuzu türleriyle kullanıldığında yetenek tahminindeki verimlilikten ödün vermeden havuzun tamamının kullanımını iyileştirmede etkili olduğu kanıtlanmıştır (Chang ve Ying, 1999). İçeriğe dayalı madde seçim yönteminde değerlendirilen belirli içerik alanıyla olan ilgisine göre maddelerin seçimine öncelik verir. Bu yaklaşım, testin içerik geçerliliği için gerekli olan ilgili içerik alanlarını yeterince örneklemesini

sağlar. Bu hedefe ulaşmak için sıklıkla içerik dengeleme algoritmaları kullanılmaktadır (Wainer, Bradlow ve Wang, 2007).

Testi bitirme kuralı ise BBT ile ilgili çalışmalarda yer alan başka bir boyuttur. İlgili alan yazında BBT oturumunu sonlandırmada kullanılan birçok kriter bulunmaktadır;

- i) Sabit madde sayısı (De Ayala, 1992),
- ii) Yeteneğin standart hatası,
- iii) Madde bilgi fonksiyonunun belirlenen bir sabit sayının altına düşmesi,
- iv) Bu kriterlerin birleştirilmesi.

Bu kriterlerden ilk ikisini dikkate alarak geliştirilen yöntemler en sık kullanılan BBT bitirme yöntemleridir (Gushta, 2003). Sabit madde sayısı ile bitirme kriterini esas alan yöntemde, birey için önceden belirlenen sayıda madde BBT oturumunda sorulmakta ve test bu belirlenen sayısı kadar madde sorulduktan sonra bitirilmektedir. Bu yöntem kapsam geçerliliğini sağlamak adına kullanışlıdır çünkü bireye ne kadar madde sorulacağı belirlidir. Ancak, bu yöntemde güvenilirliğin yüksek olacağı garanti değildir. Bireye bütün maddeler sorulduktan sonra bile BBT'nin güvenilirliğinin sağlanması için standart hata çok yüksek olabilir. Bu sorunu çözmek için BBT oturumu başlamadan önce madde güclüğü için bir standart hata belirlenebilir. Böylelikle bireylerin kestirilen yeteneklerinin güvenilir olmaları sağlanmış olur. Bu durum da testin kapsam geçerliliğini tehlikeye atabilir çünkü bazı alt alanlardan madde bireye sorulmadan test bitebilir (Choi, Grady ve Dodd, 2011).

### ***Cevaplama Süresi***

BBT uygulamaları ile ilişkili önemli kavramlardan biri de test maddelerinin cevaplanma süresidir. Cevaplama süresi; maddenin sınava girene ilk kez sunulduğu andan, bir yanıt vererek veya vermeksizin, sınava giren kişinin bir sonraki maddeye geçtiği ana kadar olan ve bu maddenin daha sonraki inceleme veya tekrar ziyaret edildiği zaman olarak ifade edilebilir (Wang, 2017). Test oluşturma sürecinde bir madde için yeterli cevaplama süresinin verilmemesi elde edilen puanların geçerliliğini olumsuz etkiler. Bu nedenle cevaplama süresinin

bilinmesi, madde ve yetenek parametre kestirimlerinin doğru hesaplanmasında olduğu kadar eğitimde test geliştirme, test geçerliği ve etkili bir sınav için de önemli bir faktördür (Altuner, 2019).

Öğrencilerin bir maddeyi cevaplamak için harcadıkları süreler farklı değişkenlerden etkilenebilir. Cevaplama süreleri bireyin yetenek düzeyi ve maddelerin güçlük düzeyinden, (Altuner, 2019; Goldhammer ve ark., 2014; İlgün Dibek, 2020; Yavuz, 2019) dikkatsizlik gibi kişisel özelliklerin yanı sıra madde türü ve içeriğinden de etkilenebilmektedir. Örneğin, ters kodlanan maddeler, normal olanlardan daha uzun cevaplama sürelerine sahip olabilir verebilir (Kensinger ve ark., 2007). Ayrıca, eğitim değerlendirmeleri bağlamında, cevaplama süreleri, çoktan seçmeli ve boşluk doldurma maddelerinden de etkilenebilir (De Boeck ve ark., 2011). (Kensinger ve ark., 2007). Hız testlerindeki zamanın yetmeme sıkıntısı veya testlerdeki fazla zaman verme sorununu ortadan kaldırılabilmek için her bir maddeye yönelik ortalama cevaplama sürelerin hesaplanması önemlidir. Diğer yandan testleri daha hızlı tamamlayabilen öğrencilerin, yavaş ilerleyen öğrencilere göre daha yüksek becerilere sahip olduğu varsayıldığı bilinmektedir. Bu varsayımla ilişkili olarak cevaplama sürelerinin hesaplanma amacı, öğrencilerin yetenek tahminlerinin doğruluğunu artırmaktır (Bagus ve ark., 2022). Ayrıca cevaplama süresi madde seçimi hakkında bilgi vererek madde performans analizlerinde katkı sağlamak ve bu süreye ilişkin verileri analiz etmek, cevaplayanlar ile maddeler arası etkileşimi anlamamıza yardımcı olabilmektedir (Ju, 2021).

Psikolojik değerlendirmelerde, tıklamalar, tuş vuruşları ve zaman damgaları gibi çevrimiçi anketlerden elde edilen meta veriler, sınava giren kişinin davranışı hakkında ek bir bilgi kaynağı sağlayabilmektedir. Bu tür verilere başka bir örnek, madde düzeyindeki cevaplama süreleridir (Paulhus, 2002). Cevaplama süresi, öğrencinin yanıtı tıklayana veya yazana kadar bir madde üzerinde harcadığı süredir. Bir öğrenci bir madde üzerinde çok az zaman harcarsa, doğru cevabı bulmak için çaba sarf etmediği sonucuna varırız (Ofek Shanny, 2020). Wise ve Kong (2005) bu tür bir tepkiyi, “*çözüm davranışının*” aksine, “*hızlı tahmin etme davranışı*” olarak adlandırmakta ve hızlı tahmin etme tepkilerinin filtrelenmesini önermektedir.

Cevaplama süresi; BBT uygulamalarında madde seçiminde, test-yanıt motivasyonu ilişkisinde, bireyin test çözme davranışını temel alarak anormal tepki davranışlarını açığa çıkarmada, kesinliği artırmada kullanıldığı gibi ayrıca ek bir bilgi kaynağı olarak bireylerin yetenek ve madde parametre kestiriminin veya öğrenci başarısının bir göstergesi olarak da kullanılabilir (van der Linden ve Guo, 2008; Wise ve DeMars, 2010; Lee ve Jia, 2014; İlgün Dibek, 2020; Yalçın, 2022).

Gelişen bilgisayar teknolojileri cevaplama süresi ölçümlerini yüksek güvenilirlikte elde edilebilir hale getirmiştir. Zorluk derecesi yüksek olan maddelerin, sınava giren kişinin daha çok zamanını aldığını belirtmek makul bir tahmin olabilir. Hâlbuki bu kişi maddeyi doğru cevaplamanın maddenin zorluğundan öte kendi yeteneği ile alakalı olduğunu fark ettiği sürece bu tahmin her zaman doğru olmayabilir. Kişiler hızlı bir şekilde cevap verdiği takdirde bir sonraki madde için daha çok zamana sahip olur. Fakat bu durum yeterli bir toplam puan elde etmek için iyi bir strateji olmayabilir. Sınava giren kişilerin BBT'ye devam ederken kullanabilmesi için bazı stratejiler önerilmektedir. Bu stratejilerin ilkinde çoğunlukla adaylara, ilk birkaç maddede emin olmak için daha çok zaman harcamaları tavsiye edilir. Bu sayede BBT'lerde adayların baştaki 5 – 10 maddeyi doğru cevapladıkları görülmüştür. Diğer bir stratejiye göre, adaylar kendilerine zor gelen maddelerde hızlı hareket etmelidirler. Bu sayede de doğru cevaplayabilecekleri maddelere daha çok vakit ve bilişsel kaynak ayırabilmektedirler. Bu stratejilerin kullanımı BBT'nin son yetenek tahminlerinde önemli bir etkiye sahiptir (Steffen and Way, 1999).

Bergstrom, Gershonve Lunz (1994) BBT'lerde kişilerin yanlış cevapladıkları maddelerde doğru cevapladığı maddelerden daha fazla zaman harcadıklarını gözlemlemiştir. Sınava girenlerin doğru veya yanlış cevapladıkları maddelere harcadıkları farklı miktardaki toplam zamanın, onlara verilen maddelerin zorluk seviyelerini yansıtabileceği varsayımı mantıklı olacaktır. Farklı maddelere harcanan süreler, test katılımcıları için maddelerin nispeten zorluk seviyelerini gösteren bir gösterge olarak hizmet edebilir. Daha zor maddeler doğru cevapları almak için daha fazla zaman gerektirebilmektedir. Schnipke ve Scrams (1997),

testi temsilen seçilen ve zorluk dereceleri kolaydan zora doğru değişen dört maddedeki madde performansına bakarak, maddelerin zorluk derecesi arttıkça aynı doğruluk düzeyine ulaşmak için gereken cevaplama süresinin arttığını bulmuşlardır. Fakat başka bir araştırmada ise zorluk düzeyi ne olursa olsun, yanlış cevaplanan maddelere yanıt vermek için daha uzun yanıt sürelerinin gerektiği belirtilmiştir (Hornke, 2000). BBT'lerde maddeler sınavı alan adaylara bağlı olduğu için, bütün adaylar maddeleri kendi kişisel yetenek seviyelerine göre 'kolay' veya 'zor' olarak nitelendireceklerdir. Bir madde yüksek yetenekli sınava giren kişi için kolayken düşük yetenekli için zor olabilir. Bu yüzden BBT'lerde, yüksek ya da düşük yetenekli kişiler açısından kolay ya da zor maddelere yönelik harcanan zaman farklılıkları azaltılabilir.

### **Lognormal Cevaplama Süresi Modeli**

Lognormal cevaplama süresi modeli, madde cevaplama süresinin modele dahil edilerek madde parametrelerinin ve yetenek parametrelerinin kestirildiği modellerdendir (LNCSM). Bu modelde, yanıtlayıcının madde cevaplama ne kadar zaman harcadığını tanımlayan bir lognormal model ile analiz yapılır. Gizil değişken olarak yetenek ile hızın birlikte incelendiği modeldir. MTK'da kullanılan 2 PLM ile benzer özellikler taşımaktadır. Her bir yanıtlayıcının hızını temsil eden parametrelerin yanı sıra bu modelde her bir test maddesinin süre yoğunluğunu ve süre ayırt ediciliğini belirten parametreler de hesaplanır (van der Linden, 2006). Süre yoğunluğu ( $\lambda$ ) maddeyi tamamlamak için gerekli olan ortalama süreyi (logaritmik ölçekte) ve süre ayırt ediciliği ( $\Phi$ ) çalışma hızının cevaplama süresi üzerindeki maddeye özgü etkisini temsil etmektedir. Hız ve madde parametrelerinden elde edilen cevaplama süresinin (response time – RT) normal dağıldığı varsayılmaktadır (Fox ve ark., 2023).

$$RT_{ik} = \lambda_k - \phi_k \zeta_i + \epsilon_{ik}$$

$$\epsilon_{ik} \sim N\left(0, \sigma_{\epsilon_k}^2\right). \quad (1)$$

Kişi ve madde parametreleri çok değişkenli modellerde açıklayıcı değişkenler de içererek genişletilebilir.  $X_\theta$  yetenek parametresini ve  $X_\zeta$  hız parametresini göstermek üzere kişi parametrelerinin ortalaması (2) numaralı eşitlikteki gibi hesaplanabilir.

$$\mu_{\theta} = \mathbf{X}_{\theta}\beta_{\theta}, \mu_{\zeta} = \mathbf{X}_{\zeta}\beta_{\zeta}. \quad (2)$$

Bu model, cevaplama doğruluğu (yetenek) ve cevaplama süresi (hız) olarak iki bileşene sahiptir. Hız düştüğü zaman doğruluk artacağı için yetenek parametresi artma eğiliminde olduğu zaman hız parametresi düşme göstermektedir (Fox ve ark., 2023).

### Problem Durumu

BBT uygulamalarında sıklıkla bireylerin maddeleri cevaplama süreleri kaydedilmektedir. Ancak, madde cevaplama sürelerinin de dahil edilerek yetenek ve parametre kestirimi yapılan matematiksel modeller simülasyon çalışmaları ile desteklenmektedir. Bu modeller gerçek veri ile çalışıldığı zaman da madde cevaplama süreleri genellikle bütün cevaplayıcılara aynı maddelerin aynı sırayla geldiği bilgisayar tabanlı testlerden elde edilmektedir.

BBT uygulamalarının kullanımının artmasına rağmen madde cevaplama süresinin madde parametresi olarak pek ele alınmaması, madde cevaplama süresini dikkate alan MTK modellerinin gerçek BBT uygulamalarında kullanımını yaygınlaştıramamaktadır. Halbuki BBT uygulamasında bireyin hızının ve madde cevaplama süresinin dikkate alınması ile hem bireye sunulacak madde sayısı hem de bireyin testte harcayacağı süre değişebilir. Bu durum bireyin yeteneğine daha uygun madde sorulmasını ve böylelikle bireyin testte daha iyi odaklanmasını sağlayabilir.

Madde cevaplama süresinin kağıt-kalem testlerinden elde edilememesinden dolayı bilgisayar tabanlı testlere ihtiyaç duyulmaktadır. Zor bir maddeye farklı yetenek düzeyindeki bireyler farklı zaman harcayacaklardır. Aynı durum kolay maddeler için de geçerlidir. BBT uygulamalarında düşük yetenek düzeyindeki bireylere zor maddeler sorulmayacağı için madde cevaplama süresini tüm katılımcılara aynı maddelerin sorulduğu bilgisayar tabanlı testler yerine BBT uygulamalarından elde etmek daha iyi olabilir. Böylelikle aynı yetenek düzeyindeki bireylerin benzer maddeleri cevaplamak için harcadıkları sürenin değişmesi yeteneğin de farklılaşmasına sebep olabilir. Bu farklılaşmaları belirleyebilmek ve tartışabilmek için BBT

uygulamalarında madde cevaplama süresini dikkate almayan MTK modellerinin yanı sıra madde cevaplama süresini dikkate alan MTK modellerine de yer verilmelidir.

### **Araştırmanın Amacı ve Önemi**

BBT uygulamalarında yüksek yetenekli bireylere zor maddeler geldiği ve bireylerin de bu maddelerde daha fazla zaman harcadığı bilinmektedir. Aynı zamanda BBT uygulamalarında maddeyi cevaplama hızı geçme özelliği sıklıkla kullanılmadığı ve bireylerin maddeye tekrar geri dönme seçenekleri olmadığı için bireyler, maddenin cevabından emin olmasalar da o maddeye cevap vermek zorundadırlar. Ayrıca BBT uygulamalarında bireyin cevaplayacağı madde sayısı sabit olmadığından ve bilinmediğinden bireyler maddeye harcadıkları süre konusunda endişeye düşüp, sıklıkla fazla zaman harcadıklarında, doğru cevaba ulaşmasalar bile maddeyi yanıtlayıp geçmektedirler.

Benzer yetenek düzeyine sahip iki bireyin benzer parametrelere sahip maddeleri cevaplama hızları da değişmektedir. Madde cevaplama süresi, BBT uygulamalarında madde parametreleri arasında çok yer almamaktadır. Maddeyi doğru ve hızlı cevaplamak da yeteneğin bir göstergesi olabilir. Bu sebeplerle, madde cevaplama süresinin yetenek ve parametre kestirimine dahil edildiği modeller özellikle BBT uygulamalarında önem kazanmaktadır. Bu araştırma, gerçek BBT uygulamasından elde edilen verilerle madde cevaplama süresinin dahil edildiği modellerden elde edilen madde parametreleri ve yetenek düzeylerini incelediği için önemlidir. Bu çalışmada, bu ilişkiyi ortaya çıkarmak amaçlanmıştır.

### **Araştırma Problemi**

Bilgisayarda bireyselleştirilmiş testlerde her bir maddenin cevaplandırılma süreleri, madde parametreleri ve yetenek düzeyleri ile ne düzeyde ilişkilidir?



## Alt Problemler

Araştırmanın amacı doğrultusunda şu alt problemlere cevap aranmıştır;

1. Yetenek düzeyleri ile bireylerin cevapladıkları madde sayıları, madde parametreleri, madde cevaplama süreleri ve test toplam süreleri arasında istatistiksel olarak anlamlı ilişki var mıdır?
2. Düşük ve yüksek yetenek düzeylerine göre bireylerin cevapladıkları madde sayıları, madde parametreleri, madde cevaplama süreleri ve test toplam süreleri arasında istatistiksel olarak anlamlı fark var mıdır?
3. Maddenin doğru veya yanlış cevaplanmasına göre madde cevaplama süresinde istatistiksel olarak anlamlı fark var mıdır?
4. Madde cevaplama süresini dikkate alan lognormal cevaplama süresi modeli (LNCSM) ile kestirilen madde parametreleri ve madde cevaplama süresini dikkate almayan MTK modelleri ile kestirilen madde parametreleri arasında istatistiksel olarak anlamlı ilişki var mıdır?
5. Madde cevaplama süresini dikkate alan lognormal cevaplama süresi modeli (LNCSM) ile kestirilen yetenek düzeyleri ve madde cevaplama süresini dikkate almayan MTK modelleri ile kestirilen yetenek düzeyleri arasında istatistiksel olarak anlamlı ilişki var mıdır?

## Sayıtlılar

Araştırmada, katılımcıların bilgisayar kullanım becerilerinin sonuçlara yansımadağı ve BBT’de yer alan maddeleri, gerekli süreyi harcayarak cevapladıkları varsayılmıştır.

## Sınırlılıklar

Araştırmada Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı’na (ALES) ait veriler sınırlı süresi olan bir kağıt-kalem uygulamasından toplanmıştır. ALES, sayısal ve sözel

yeteneđi ölçen bir sınavdır. Her test için ayrı ayrı deđil toplam süre vardır ancak sınava giren adayların belirli bir süreleri olduđu için ALES'in hız testi olabilmesi durumu bu araştırmanın sınırlılıklarındandır.

Araştırmada geliştirilen BBT uygulamasının nihai bir süre sınırı bulunmamaktadır ancak adayların maddeleri cevaplarırken harcadıkları süre önemli olduđu için madde bazında süre sınırlılığı vardır. Ayrıca hazırlanan BBT uygulamasında madde seçimi ve yetenek kestirimi yapılırken madde cevaplama süresinin dikkate alınmaması da bu araştırmanın sınırlılıklarındandır.

Bu araştırma kađıt-kalem uygulamasında kullanılan MTK yöntemi (Rasch Model), BBT uygulamasında kullanılan MTK yöntemleri (Rasch Model, tabakalı örnekleme, BM ve MFI) ve BBT'den elde edilen süre parametresinin kullanıldıđı LNCSM ile sınırlıdır.

## **Tanımlar**

CONCERTO: Cambridge Üniversitesi Psikometri Merkezi tarafından geliştirilen, eğitim ortamlarında bilgisayarda bireyselleştirilmiş testlerin oluşturulması ve yönetilmesini desteklemek üzere tasarlanmış web tabanlı bir platformdur.

## Bölüm 2

### Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

#### Bilgisayarda Bireyselleştirilmiş Testler İle İlgili Yapılmış Çalışmalar

Kalender (2011), Öğrenci Seçme Sınavı'nın (ÖSS) fen alt testinin kağıt-kalem formu ile BBT formundan elde edilen sonuçları karşılaştırmıştır. Bu karşılaştırmayı yaparken farklı yetenek kestirim yöntemlerini ve farklı test sonlandırma kurallarını dikkate almıştır. Maximum Likelihood Estimation ve Expected A Posteriori yetenek kestirim yöntemlerini ve standart hata eşik değeri ile sabit madde sayısı test sonlandırma kurallarını kullandığı çalışma iki aşamalıdır. Çalışmanın ilk aşamasında kağıt-kalem ve BBT formatından elde edilen yetenek kestirim yöntemlerini karşılaştırmak için post-hoc simülasyon yapmıştır. Yetenek kestirim yöntemi olarak Expected A Posteriori'nin ve testi sonlandırma kuralının 0,30 ya da daha yüksek standart hata eşik değerinin uygun olduğu simülasyon ile belirlendikten sonra BBT uygulaması hazırlamıştır. 33 öğrencinin katıldığı BBT uygulamasında uygulanan madde sayısı ortalaması 18,4 olmuştur. BBT ve ÖSS'nin kağıt-kalem formu ile elde edilen yetenek kestirimleri arasındaki ilişkiyi ise 0,74 olarak hesaplamıştır. Yaklaşık %50 oranında daha az sayıda madde ile daha yüksek güvenirliliğe sahip yetenek kestirimi BTT uygulaması ile elde etmiştir.

Kezer (2013), MTK'nın bir uygulaması olan BBT yöntemi ile geleneksel kağıt-kalem test yönteminin karşılaştırılmasını ve BBT yöntemine ilişkin farklı stratejilerin değerlendirilmesini amaçlamıştır. Ankara Üniversitesi Yabancı Diller Yüksekokulu, 2012-2013 eğitim öğretim yılında hazırlık sınıfında eğitim alan toplam 1166 öğrenciden elde ettiği İngilizce Kelime Testi verileriyle bir araştırma tasarlamıştır. R açık kaynaklı istatistik programının "catR" kütüphanesini farklı stratejileri karşılaştırmak için kullanmıştır. Başlatma kurallarını, yetenek kestirim yöntemlerini ve sonlandırma kurallarını değiştirerek simülatif veriler oluşturmuştur. BBT uygulaması ile, kağıt-kalem testlerine göre daha az madde kullanıldığı sonucuna ulaşmıştır. Diğer yandan, BBT'de kullandığı farklı stratejiler ve kağıt-kalem testinden elde ettiği yetenek parametreleri arasında pozitif yönde yüksek korelasyon katsayıları bulmuş, böylece

araştırma kapsamında ele aldığı 18 farklı strateji ile kağıt-kalem testlerinin çok benzer yetenek parametreleri kestirdiğini ortaya koymuştur. Buna ek olarak, yetenek kestirim yöntemini değiştirse de elde edilen yetenek parametreleri arasında pozitif ve yüksek korelasyon olduğunu göstermiştir. Sonlandırma kurallarını dikkate aldığına, farklı stratejilerden elde edilen yetenek kestirimlerinin, gerek kağıt-kalem testinden elde edilen yetenek parametreleri ile gerekse kendi aralarında, en düşük korelasyon katsayılarına standart hatanın 0,50'den küçük olduğu durumlarda ulaşıldığını gözlemlemiştir. ML, EAP ve MAP yetenek kestirim yöntemlerinden kaynaklı, kestirilen yetenek parametrelerinde farklılık olmadığını belirlemiştir. Ayrıca, ML yönteminin EAP ve MAP yöntemlerine göre yetenek kestirirken daha fazla madde kullandığı sonucuna ulaşmıştır.

Özbaşı (2014), Ankara Üniversite'sinde birinci sınıf öğrencilerine uygulanan "*Bilgi ve İletişim Teknolojileri*" dersi için yapılan muafiyet sınavı kapsamında öğrencilerin yeterliliklerini ölçmeye yönelik uygulanan BBT olarak kullanılabilirliğini araştırmıştır. İki aşamalı olarak yürütülen araştırmada ilk olarak 1366 üniversite birinci sınıf öğrencisinden elde edilen kağıt-kalem verileri ile simülatif BBT uygulaması geliştirmiştir. En Yüksek Olabilirlik Yaklaşımı (EYOY) ve Beklenen Sonsal Dağılım (BSD) yetenek kestirim yöntemleri ile sabit madde ve standart hata ( $SH < 0,30$  ve  $SH < 0,50$ ) test sonlandırma kuralları simülasyonda kullanmıştır. EYOY yetenek kestirim yönteminin ve sabit madde ile standart hatanın ( $SH < 0,30$ ) test sonlandırma kuralı olarak kullanıldığı, 142 öğrencinin katıldığı gerçek BBT uygulaması araştırmanın ikinci aşamasını oluşturmaktadır. Araştırmacı, geliştirdiği "*Bilgisayar Okur Yazarlık Testi*" BBT uygulamasında kullanmış ve güvenilir bir şekilde yeterli ölçülerinin kağıt-kalem testine göre daha az madde ile elde edilebileceğini göstermiştir.

De Beurs ve ark. (2014), Beck İntihar Düşüncesi Ölçeğinin uzunluğunu azaltmak için MTK'ya dayalı bir BBT simülasyonu çalışması yapmışlardır. Araştırmada simülasyon için kullanılan veriler, Hollanda'da yapılan "*İntiharı DURDURMAK için Eğitim Alan Profesyoneller (PITSTOP suicide)*" çalışmasından elde etmişlerdir. Veri setine temel bileşenler analizi, doğrulayıcı faktör analizi, kademeli yanıt modeli (GRM) uygulanmış ve BBT uygulaması

geliştirmişlerdir. Geliştirilen uygulamada Beck İntihar Düşüncesi Ölçeğinin 19 maddelik tam halini 505 kişiye uygulamışlardır. Uygulama sonucu yapılan psikometrik analizler testin yüksek iç tutarlık ile tek boyutlu yapıda olduğunu göstermiştir. BBT uygulaması ile gelecekteki intihar davranışı riskinin artmasının tahmini için 19 yerine ortalama 4 maddenin yeterli olduğu sonucuna ulaşmışlardır.

Aybek (2016), öğrenci yönlendirme hizmetlerinde sıklıkla kullanılan Kendini Değerlendirme Envanteri'nin (KDE) BBT olarak uygulanabilirliğini araştırmak amacıyla bir araştırma yapmıştır. 1144 lise öğrencisinden elde ettiği verilere dayanarak post-hoc simülasyonlar gerçekleştirmiştir. Simülasyonlar sırasında 0,30; 0,40 ve 0,50 standart hata durdurma kuralları, Genelleştirilmiş Kısmi Puan Modeli (GKPM) ve Kademeli Tepki Modeli (KTM) gibi madde tepki modelleri, ayrıca Fisher En Yüksek Bilgi (FEYB), Beklenen En Yüksek Bilgi (BEYB), Beklenen En Düşük Sonsal Varyans (BEDSV) ve Beklenen En Yüksek Sonsal Ağırlıklandırılmış Bilgi (BEYSAB) gibi madde seçim yöntemleri kullanmıştır. Yaptığı simülasyon sonuçlarına göre, KDE'nin 23 farklı faktörü için ayrı ayrı değerlendirmeler yapmış ve BBT uygulamasında en uygun MTK modeli olarak Genelleştirilmiş Kısmi Puan Modeli'ni (GKPM), madde seçim yöntemi olarak Fisher En Yüksek Bilgi'yi (FEYB), ve sonlandırma kuralı olarak da 0,40 standart hatasını tercih etmiştir. Ayrıca, kağıt-kalem formunda 230 madde içeren KDE'nin, simülasyon sonuçlarına göre 113 madde ile sonuçlandırılmış olduğu sonucu elde etmiş ve kağıt-kalem formundan alınan puanlar ile simülasyonla kestirilen  $\theta$  düzeyleri arasında 0,90 ile 0,96 arasında korelasyon katsayıları hesaplamıştır. Post-hoc simülasyon sonuçlarına dayanarak canlı BBT uygulamasını geliştirmiş ve 25 öğrenciye uygulamıştır. Öğrencilerin kağıt-kalem formu ve canlı BBT formundan elde ettikleri profilleri karşılaştırılmış ve benzer profillerin elde edildiğini tespit etmiştir. Bunun yanı sıra, öğrencilerin kağıt-kalem formundan aldıkları puanlar ile canlı BBT formu ile kestirilen  $\theta$  düzeyleri arasındaki korelasyonların üç faktör için 0,45 ile 0,55 arasında; yirmi faktör için ise 0,63 ile 0,88 arasında değişmekte olduğunu göstermiştir. Böylelikle, KDE'nin BBT olarak uygulanabilir olduğu sonucuna bu araştırma ile varmıştır. Ayrıca, madde havuzunun daha yüksek bilgi sağlayan

maddeler ile genişletilmesinin, BBT envanterinin daha elverişli hale getirilmesine yardımcı olabileceğini de öneri olarak sunmuştur.

Şahin (2017), Anadolu Üniversitesi Yabancı Diller Yüksekokulu'nda eğitim gören hazırlık sınıfı öğrencilerinin lisans eğitimi öncesinde vermeleri gereken Yeterlilik Tespit Sınavı'nın (YTS) verileri kullanılarak gramer, kelime bilgisi ve genel yetenekleri ölçen çok boyutlu bir BBT geliştirmiştir. Bu amaçla öncelikle, geçmiş yıllarda uygulanan YTS'lerde kullanılan dört farklı madde setinden oluşan verileri kullanılarak bir madde havuzu oluşturmuştur. Analizler sonucunda tüm madde setleri için en uygun modelin üç boyutlu bifaktör model olduğu belirlemiştir. Madde havuzunu oluşturduktan sonra, gerçek-zamanlı çok boyutlu BBT uygulamasında kullanılacak algoritmayı belirlemek için eksik veri matrisini de göz önünde bulundurularak bir hibrit simülasyon yapmıştır. Bu simülasyonda farklı yetenek kestirim yöntemleri, madde seçim yöntemleri ve sonlandırma kuralları kullanılarak toplam 36 farklı koşul yaratmıştır. Korelasyon ve hata istatistiklerini değerlendirerek değişken madde sayısına dayalı sonlandırma kuralının kullanıldığı koşullarda yanıtlanan madde sayısını da raporlamıştır. Elde ettiği sonuçlar, D-rule madde seçim yöntemi ve MAP yetenek kestirim yönteminin, üç farklı sonlandırma kuralı için en iyi istatistikleri verdiğini göstermiştir. Bu sebeple, gerçek-zamanlı çok boyutlu BBT uygulamasında D-rule madde seçim yöntemini, MAP yetenek kestirim yöntemini ve sonlandırma kuralı olarak standart hataya dayalı yöntemi tercih etmiştir. Daha sonra, bu algoritmayı kullanılarak hazırladığı gerçek-zamanlı çok boyutlu BBT uygulamasının sonuçları ile kağıt-kalem testiyle elde ettiği genel yetenek düzeyleri arasındaki korelasyonu hesaplamıştır. Ayrıca, uygulamaya katılan 99 kişinin sonuçlarına dayanarak havuzda yer alan maddelerin kullanım sıklıklarını ve her bir bireyin yanıtladığı toplam madde sayısını da raporlamıştır. 200 maddeden oluşan havuzun yaklaşık %30'unun kullanıldığı ve testi bitiren katılımcıların ortalama olarak 12,3 madde yanıtladığı sonucuna ulaşmıştır. Ayrıca katılımcıların yaklaşık %75'inin 10 ile 12 arasında madde yanıtladıklarını belirlemiştir.

Şenel (2017), görme engelli öğrencilerin okuyucu-kodlayıcı yardımıyla aldıkları kağıt-kalem testlerinde ve BBT uygulamasında ölçülen dinlediğini anlama becerinin karşılaştırılması ve öğrencilerin bu uygulamalara yönelik görüşlerinin tespit edilmesi amacıyla 7. ve 8. sınıfa devam eden 51 görme engelli öğrenciyle bir çalışma yapmıştır. 608 kişilik deneme grubuna 166 madde uygulamış ve bu maddelerden 142'sini BBT havuzuna alırken, 24'ü ile kağıt-kalem testi hazırlamıştır. BBT uygulamasını sesli okuma düzenlemesi biçiminde tasarlamıştır. Uygulamaya katılan 9 öğrenciyle görüşme yapmış ve uygulamaya yönelik görüşlerini belirlemiştir. BBT uygulaması sonucunda kağıt-kalem testine göre öğrencilerin yetenek puanları anlamlı bir şekilde düşük olsa da her iki testten de elde edilen yetenek puanları arasında yüksek düzeyde korelasyon elde edildiği için benzer yetenek kestirimlerin yapıldığı şekilde yorumlanmıştır. Aynı zamanda BBT'nin daha az madde ile daha kısa sürede daha güvenilir yetenek kestirimi yaptığı sonucuna ulaşmıştır. Araştırmada yer verilen öğrenci görüşlerine göre öğrencilerin BBT hakkında olumlu görüşleri olsa da her katılımcıya farklı sayıda ve farklı madde gelmesi öğrencilerin testin geçerliği hakkında sorularının olduğunu göstermektedir. Öğrenciler genellikle testin uzunluğu, teste harcanan süre, maddelerin okunurken anlaşılır olması ve uygulamanın adil bir şekilde yapılması gibi ifadeler üzerine yoğunlaşmışlardır. Bunun yanı sıra, BBT ile maddeyi tekrar dinleme ve istedikleri yerde durdurma özelliğinin onlar için en işlevsel yönü olduğuna değinmişlerdir. Araştırmacı bu yorumlar eşliğinde öğrencilerin motivasyonlarının arttığını gözlemlemiştir.

Choi ve McClenen (2020), yaptıkları araştırmada öğrenme analitiği olarak Bayes ağlarını ve BBT'yi temel alan yapay zeka yöntemlerini kullanarak bilgisayarlı biçimlendirici uyarlamalı test (CAFT) adını verdikleri uyarlamalı biçimlendirici değerlendirme sistemi geliştirmeyi amaçlamışlardır. CAFT, öğrencinin yeteneğine uyumlu bir şekilde öğrenciye uygun maddeleri ve testleri seçerek bireyselleştirilmiş biçimlendirici değerlendirmeyi uyarlamalı olarak uygulayabilir. Araştırmada bir madde havuzundaki 40 madde 410 öğrenci tarafından değerlendirilmiştir. Ayrıca 1000 öğrenci bir simülasyon çalışması için alınmış ve 120 öğrenci CAFT'nin etkinliğini, geçerliliğini ve güvenilirliğini değerlendirmek için bir uygulama

çalışmasına kaydolmuştur. Elde edilen sonuçlar CAFT aracılığıyla öğrencilerin öğrenmelerindeki ilerlemeleri hakkında kişiselleştirilmiş geri bildirim alabildiklerini göstermiştir. Araştırmacılar, sonuç olarak, yapay zeka bileşeni olarak BBT'yi entegre etmiş bir öğrenme yönetim sisteminin bireyselleştirilmiş uzaktan öğrenme hizmetleri için etkili bir değerlendirme aracı olabileceğini vurgulamışlardır.

Gür ve Gülleroğlu (2020), 1-0 şeklinde puanlanan maddelerin BBT uygulamalarında, madde kullanım sıklığı kontrol yöntemlerinin ölçme duyarlılığına ve test güvenliğine etkisini incelemek için bir çalışma gerçekleştirmiştir. Çalışma, farklı örneklem büyüklükleri, test uzunlukları ve yetenek dağılımlarının etkilerini araştırmayı amaçlamaktadır. Monte Carlo simülasyonu kullanılarak yapılan bu çalışmada, 50 maddeden oluşan uzun bir test ve 35 maddeden oluşan kısa bir test tasarlanmıştır. Büyük grup örnekleminde 1000 kişi ve küçük grup örnekleminde 250 kişi yer almıştır. Testi alan bireylerin yetenek parametresi ( $\theta$ ) değerleri -3 ile +3 aralığında sola çarpık, tek biçimli, normal ve sağa çarpık dağılımlar şeklinde oluşturulmuştur. "Simpson Hetter Stratejisi" ve "Aşamalı Düşürme" adlı iki madde kullanım sıklığı kontrol yöntemi kullanılmıştır. Ayrıca, madde kullanım sıklığının kontrol edilmediği bir referans koşulu da belirlenmiştir. Bu durumda toplamda 48 simülasyon koşulu oluşturulmuş ve her bir koşul için 100 tekrarlamayla toplamda 4800 veri dosyası elde edilmiştir. Her bir koşul için yapılan ölçme duyarlılığı ve test güvenliği hesaplamaları sonucunda, madde kullanım sıklığı kontrolü koşullarında genel olarak ölçme duyarlılıklarında büyük farklılıklar bulunmamıştır. Ancak Aşamalı Düşürme yöntemi seçildiğinde, test güvenilirliği daha iyi sağlanmıştır. Sonuç olarak, Aşamalı Düşürme yöntemi seçildiğinde, madde kullanım sıklığının dengeli bir şekilde kontrol edilerek madde havuzunun daha verimli bir şekilde kullanıldığı ve test güvenliğinin daha iyi sağlandığı sonucuna ulaşılmıştır. Ancak, ölçme duyarlılığının düşürülmediği bu durumda, Simpson Hetter stratejisi BBT uygulamalarında en yaygın kullanılan madde kullanım sıklığı kontrol yöntemidir.

Nguyet ve Daud (2021), Güney Doğu Asya'da Vietnam kökenli bir öncü proje geliştirmek amacıyla BBT oluşturmuşlardır. Bu çalışmada, akademik performans, aile, öğretim



görevlisi ve finans olmak üzere beş alt alana dayalı 68 maddelik bir madde havuzu tasarlanmış ve 2085 öğrenci (704 erkek ve 1381 kadın) üzerinde değerlendirme yapılmıştır. Çok Boyutlu Rastgele Katsayılar Çok Terimli Logit (MRCML) Model, Çok Boyutlu Stres Ölçeklerine, BBT geliştirmek için uygulanmıştır. Elde edilen sonuçlar, MRCML'nin psikometrik özelliklere sahip yeni bir ölçek geliştirmek için etkili bir araç olduğunu göstermiştir. Araştırmada yapılan Simülatif BBT ve kağıt-kalem testine ilişkin yeterlik ve güvenilirlik ölçümleri incelendiğinde, değerlerin birbirine yakın olduğu sonucuna varılmıştır. Bu da MRCML Modeli tabanlı BBT'nin, geleneksel kağıt-kalem testleriyle benzer güvenilirlik ve yeterlik düzeylerine sahip olduğunu göstermektedir.

Şahin Kürşad (2023) farklı madde seçim yöntemlerinin BBT'de Test Bilgi Fonksiyonu (TBF) ve test verimliliği üzerindeki etkisini incelemiştir. Bu çalışma simüle edilmiş verilerle gerçekleştirilmiş olup, "örneklem büyüklüğü, yetenek parametresi dağılımı, madde havuzu boyutu, MTK modeli ve madde parametrelerinin dağılımı, yetenek tahmin yöntemi, başlama kuralı, madde kullanım sıklığı kontrolü ve durdurma kuralı" gibi bazı değişkenler sabit tutulmuştur. Çalışmada araştırılan bağımsız değişkenler arasında "aralık bilgisi kriteri, etkinlik dengeli bilgi, -b değeri eşleştirme, Kullback-Leibler bilgisi, maksimum Fisher bilgisi, olabilirlik ağırlıklı bilgi ve rastgele seçim" yer almaktadır. Araştırma sonuçlarına göre, TBF açısından en iyi performansı maksimum Fisher yöntemi sağlamıştır. Ayrıca hem TBF hem de test verimliliği açısından en kötü performansa sahip olan rastgele seçim yöntemi dışında, diğer yöntemlerin performanslarının benzer olduğu belirtilmiştir.

### **Testte Cevaplama Süresi İle İlgili Araştırmalara Genel Bir Bakış**

Wise ve DeMars (2006) tarafından yapılan bir çalışmada, madde cevaplama süresini yetenek tahmini ve madde parametresi tahminine dahil eden eforla yönetilen MTK (effort-moderated IRT) modeli tanıtılmıştır. Araştırmacılar, iki çalışmada gerçek ve simüle edilmiş verilere dayanan eforla yönetilen modelin, standart 3PL modelinden daha iyi performans gösterdiğini belirtmişlerdir. Eforla yönetilen modelin daha iyi model uyumu, daha doğru madde

parametre tahminleri, daha doğru tahmin edilen test bilgileri ve daha yüksek yakınsak geçerlilik sağladığı belirtilmiştir.

Fox ve ark. (2007) çalışmasında, yanıtların ve cevap sürelerinin modellenmesine olanak sağlayan bir R paketi tanıtmıştır. Bilgisayar tabanlı testlerde yanıtların ve yanıt sürelerinin ilişkisinin karmaşık olduğu ve gözlem seviyelerine göre değişebileceği belirtilmiştir. Araştırmacılar, birleşik bir model önererek madde yanıtlarını iki parametrelili normal MTK modeliyle ve yanıt sürelerini lognormal modeliyle modellendirmişlerdir. Bu modelde, tüm parametrelerin MCMC tahmin yaklaşımı kullanılarak eş zamanlı olarak tahmin edilebildiği belirtilmiş ve MCMC algoritması için R paketi çalışmada açıklanmıştır.

Van der Linden (2008), yanıtlar ve yanıt süreleri için ayrı birinci düzey modellere sahip hiyerarşik bir modelleme çerçevesi ve sınava girenler popülasyonundaki yetenek ve hız parametrelerinin dağılımı için ikinci düzey bir model kullandığı bir çalışma yapmıştır. Oluşturulan sistemde, her yeni yanıt süresi oluşumunda yetenek parametresi için ampirik bir önceki dağılımın güçlendirilmesine izin verilmiştir. Bu çalışmada, eforla yönetilen modelin madde seçimini iyileştirmek için kullanılabileceği belirtilmiştir.

Suh (2010) çalışmasında MTK'ya dahil edilen ayırt edici yanıt süresi modellerini karşılaştırmış ve gerçek ve simüle edilmiş veriler kullanarak madde özellikleri ile sınava giren kişinin yeteneği ve yanıt süresi arasındaki ilişkiyi incelemiştir. Çalışmada, gerçek veriler üzerinde cevaplama süresinin araştırılması için Thissen'in (1983) lognormal cevap süresi modeli, Wang ve Hanson'ın (2005) 4PL RT modeli ve van der Linden'in (2007) hiyerarşik çerçevesi için Markov Chain Monte Carlo (MCMC) yöntemlerini kullanılarak Bayes tahmini uygulanmıştır. Suh (2010), gerçek veri analizinden elde edilen sonuçların van der Linden'in (2007) hiyerarşik çerçevesinin en uygun sonuçları verdiğini belirtmiştir. Simüle edilmiş veri analizinde ise van der Linden'in (2007) hiyerarşik çerçevesinin Wang ve Hanson'ın (2005) 4PL RT modelinden daha iyi sonuçlar verdiği görülmüştür.

Altuner (2019), farklı yetenek düzeylerine sahip 212 öğrenci ile madde istatistikleri ve madde cevaplama süresi arasındaki ilişkiyi incelemek amacıyla çalışma yapmıştır. Bu

çalışmaya dahil edilen öğrencilere, çoktan seçmeli İngilizce kelime bilgisi testi TAO isimli çevrim içi bir sınav platformunda uygulanmıştır. Araştırmacılar, elde edilen sonuçlara göre öğrencilerin maddelerin güçlük düzeyine göre cevaplama sürelerinin değiştiğini ve üst yetenek grubundaki öğrencilerin zor maddelere daha fazla zaman ayırıp kolay maddeleri ise daha hızlı bir şekilde cevapladığını belirtmişlerdir. Ayrıca, yapılan analizler sonucunda, madde güçlük indeksinin cevaplama süresini etkilediği, ancak madde ayırt edicilik indeksinin cevaplama süresi üzerinde anlamlı bir etkisinin olmadığı tespit edilmiştir.

Türkoğuz (2020), cevaplama süresine dayalı üç aşamalı tanılama (Three-Tier Diagnostic) ve seçmeli testlerin eşik değerlerinin karşılaştırılmasına yönelik bir çalışma yapmıştır. Çalışma kapsamında, madde cevaplama süresi doğruluk puanları (Item Response Time Fidelity Scores - RTFs) ve Cronbach's Alpha Güvenilirlik ( $\alpha$ ) katsayıları incelenmiştir. Ayrıca, 1 ile 30 saniye arasında 30 eşik noktası için KR20 katsayıları RTF'ler ile hesaplanmış ve bu değerler için grafikler oluşturulmuştur. Bu grafikler, literatürdeki eşik noktasını tahmin eden ve başlangıç verileri ile eşik değeri tarafından belirlenen yeni veriler üzerinde Cox Oransal Tehlike Modeli (Cox-HM) ile risk analizleri yapılarak desteklenmiştir. Çalışma, bir devlet üniversitesinde öğrenim gören 115 fen bilimleri öğretmen adayı ile iki grup şeklinde gerçekleştirilmiştir. 63 maddeden oluşan testte Grup-1'e (n=57) Üç Aşamalı Tanı Testi (3TDT), Grup-2'ye (n=58) ise Çoktan Seçmeli Test (MCT) uygulanmıştır. Türkoğuz (2020) tarafından elde edilen sonuçlar, cevaplama sürelerine ilişkin verilerin ölçme araçlarının geçerlik ve güvenilirlik çalışmalarında kullanılabileceğini vurgulamıştır.

Michaelides ve Ivanova (2022), düşük riskli değerlendirmelerde, testi alan bireyler testi cevaplamak için yeterli çaba göstermediklerinde, test puanları bireyin gerçek yeteneğini düşük tahmin eder ve test çözme çabasının etkisini göz ardı etmek, test sonuçlarının geçerliğine zarar verebileceğinden, Uluslararası Öğrenci Değerlendirme Programı (PISA) kapsamında ülkeler ve farklı madde türleri arasında öğrencilerin test çözme çabasının düzeyini ve doğruluğunu incelemişlerdir. 2015 PISA bilgisayar tabanlı değerlendirmesi, 59 değerlendirme alanında uygulamışlardır. Öğrencilerin test yapma çabasının davranışsal ölçüleri, Matematik ve Okuma

değerlendirmeleri için madde cevaplama sürelerine sabit ve normatif bir eşik uygulayarak hızlı tahmin etmeyi belirlemek için oluşturmuşlardır. Her bir madde üzerinde hızlı tahmin yapanların oranı ortalama olarak yaklaşık %3 olup, normatif eşikle %1 olarak bulunmuştur. İnsan kodlu açık uçlu maddelerde hızlı tahmin, basit ve karmaşık çoktan seçmeli maddelere ve bilgisayar puanlı açık uçlu maddelere göre yaklaşık iki kat daha yüksek olarak kaydedilmiştir. Hızlı tahmin yapanların ortalama performansı, tüm madde türleri için çözüm davranışı sergileyen test katılımcılarından ortalama olarak çok daha düşük ve okumada matematiğe kıyasla daha belirgin olarak bulunmuştur. Ükelere göre ağırlıklı cevaplama süreleri oldukça yüksek hesaplanmıştır ve ülke ortalaması PISA puanı ile pozitif korelasyon elde edilmiştir. Ülke düzeyinde cevaplama süresi çabasıyla diğer güçlü ilişkiler bulamamışlardır. Sonuç olarak, bilgisayar tabanlı test uygulamaları, cevaplama süresini öğrencinin test yapma çabasının bir göstergesi olarak kullanmayı kolaylaştırdığı için performans karşılaştırmalarından önce bu davranışı izleyerek ülkeler arasındaki farkların belirlenebileceği ve değerlendirmeye katılımı teşvik etmek için müdahaleler geliştirilebileceği sonucuna ulaşmışlardır.

Yalçın (2022), Türkiye'de 8. sınıf düzeyinde Trends in International Mathematics and Science Study (TIMSS) 2019 çalışmasına katılan Türk öğrencilerin fen ve matematik sorularını cevaplamak için harcadıkları sürenin yeterlik düzeyleri, özgüvenleri ve madde özelliklerine göre anlamlı bir farklılık gösterip göstermediğini incelemek amacıyla çalışma yapmıştır. Çalışmanın katılımcıları, Kitapçık 1 ve 2'deki ortak 24 (11 matematik ve 13 fen) maddesini cevaplayan toplam 577 öğrenciden oluşmaktadır. Elde edilen bulgulara göre, madde türü ve bilişsel düzeyin öğrencilerin madde cevaplama süreleri üzerinde anlamlı bir ilişkiye sahip olduğu belirlenmiştir. Ayrıca, öğrencilerin açık uçlu maddelere göre çoktan seçmeli maddelere daha fazla zaman harcadığı gözlemlenmiştir. Uygulama düzeyindeki maddeler için harcanan sürenin ise bilgi düzeyine göre anlamlı olarak daha yüksek olduğu tespit edilmiştir. Ancak, uygulama düzeyinde ve muhakeme düzeyinde maddeleri yanıtlamak için harcanan süre arasında anlamlı bir fark bulunmamıştır. Araştırmada ayrıca, öğrencilerin fen bilimlerine olan güven düzeyleri yüksek olduğunda doğru cevaplama oranlarının da yüksek olduğu ve

maddeleri daha kısa sürede cevapladıkları belirtilmiştir. Matematikte biraz özgüven sahibi olan öğrencilerin ise zor matematik maddelerinde daha başarılı oldukları ve bu tür soruları cevaplamak için daha az zaman harcadıkları ifade edilmiştir.

Aydoğan (2022), PISA 2015 uygulamasına katılan 42 farklı ülkeden 5232 öğrenci ile öğrencilerin okuma becerilerine yönelik yetenekleri ve hızlarına dair kişisel ve materyal temelli parametrelerin tahmin edilmesi hakkında bir araştırma tasarlamıştır. Araştırmada PISA 2015 uygulamasında kullanılan 36. kitapçıktaki okuma becerileriyle ilgili soruları aynı sırayla yanıtlayan öğrencilerin verilerini kullanmıştır. Doğruluk ve süre düzeylerini etkileyen faktörlerin belirlenmesi için regresyon tekniği ile analiz yapmıştır. Genetik algoritmalar yöntemi ile bu faktörleri belirlemiş, daha sonra yetenek ve hız parametrelerini açıklayıcı değişkenler olarak hiyerarşik modele eklemiştir. Daha sonra cevap süreleri ve açıklayıcı değişkenlere ilişkin veriler ile hiyerarşik modelin çözümlenmesini yapmıştır. Son olarak, bireylerin hız ve yeteneklerini yansıtan davranışların uyum düzeylerini kişi uyum istatistiklerini kullanarak ele almıştır. Sonuç olarak, hız ve yetenek parametreleri arasında zayıf düzeyde negatif bir ilişki olduğunu elde etmiştir.

Tancoš ve arkadaşları (2023), 514 çocukla yaptıkları çalışmada, daha yüksek yetenek düzeyine sahip çocukların yanlış cevap vermelerinin daha uzun sürüp sürmediğini incelemiştir. Çalışmada 29 görevden oluşan bir oyunda, basit toplama, çıkarma, çarpma ve bölme işlemleri yer almış ve cevaplama süreleri incelenmiştir. Araştırmacılar, görevlerin yanlış çözüldüğü durumların genellikle orta ve çok zor maddelerle sınırlı olduğunu ve yetenek düzeylerinin bir görevi çözmek için gereken süreyi tahmin ettiğini belirtmişlerdir. Araştırmada ayrıca, yetenek düzeyi yüksek olan çocukların maddeleri yanlış cevaplama süresinin daha uzun sürdüğü gözlemlenmiştir. Aynı yetenek düzeyine sahip çocukların ise çok kolay veya zor görevlere kıyasla görevleri çözmek için daha fazla zaman harcadıkları belirtilmiştir. Yetenek, görev zorluğu ve cevap doğruluğu arasındaki ilişkinin karmaşık olduğu vurgulanmış ve eğitim profesyonellerinin mesleki değerlendirmelerini öğrencilerin çabukluğuna dayandırmamalarının önemi vurgulanmıştır. Hızlı olmanın akıllı olmakla eşdeğer olmadığını altı çizilmiştir.

## **Bölüm 3**

### **Yöntem**

Bu bölümde araştırmanın türü, katılımcılar, veri toplama süreci, veri toplama araçları ve verilerin analizi kısımları yer almaktadır.

#### **Araştırmanın Türü**

Bu araştırmada, bilgisayarda bireyselleştirilmiş testlerde madde cevaplama süresinin yetenek düzeyi ve madde parametreleri ile olan ilişkisini belirlemek amaçlanmıştır. Bu amaç doğrultusunda ALES'in 2017 yılı sonbahar dönemi ve 2018 yılı ALES/1-2-3 oturumlarında kullanılan maddeler ve bu maddelere ait parametreler kullanılarak gerçek zamanlı BBT hazırlanmıştır. Hazırlanan bu test, ALES'e girme durumunda olan üniversite son sınıf öğrencilerine ve lisans mezunlarına uygulanmıştır. Uygulama sırasında öğrencilerin maddeleri cevaplarken harcadıkları süreler kaydedilmiştir. Elde edilen zaman parametresi yani madde cevaplama süresi kullanılarak lognormal cevaplama süresi modeli ile madde parametreleri ve yetenek düzeyleri kestirilmiş ve daha önceden ALES'in kağıt-kalem uygulamasından elde edilen madde parametreler ve yetenek düzeyleri arasındaki ilişki incelenmiştir.

İki veya daha fazla ölçülebilir veri arasındaki ilişkiyi ve onun derecesini açığa çıkarmak için gerekli veriyi toplama sürecini de içeren çalışmalara ilişkiyel araştırma denir. İlişkiyel araştırmaların amacı veriler arasındaki ilişkiyi incelemek veya bu ilişkileri tahmin yapmak için kullanmaktır (Mills ve Gay, 2019). Bu sebeple bu araştırmanın türü ilişkiyel araştırma olarak ele alınabilir.

#### **Çalışma Grubu**

Gerçek zamanlı BBT'yi hazırlamak için kullanılan madde parametreleri ALES'in kağıt-kalem uygulamasından elde edilmiştir. ALES'e 2017 yılının sonbahar döneminde 221894, 2018 yılının ALES-1 dönemine 272866, 2018 yılının ALES-2 dönemine 119064 ve 2018 yılının

ALES-3 dönemine 116174 kişi katılmıştır. BBT için kullanılacak madde parametreleri katılımcıların tamamının cevapları göz önüne alınarak hesaplanmıştır.

Bu araştırmada ALES maddeleri ile hazırlanan ölçme aracı uygulandığı için araştırmaya lisans son sınıf öğrencileri ile lisans mezunu öğrencilerin katılması uygundur. Bu sebeple gerçek zamanlı geliştirilen BBT bu özellikleri taşıyan kişilere ulaştırılmaya çalışılmıştır. Geliştirilen BBT'nin sayısal testine 365 kişi başlamış ancak 238 kişi testi tamamlamıştır. Sözel testine ise 219 kişi başlamış, 203 kişi testi tamamlamıştır. Sayısal testle ilgili analizler 238, sözel test ile ilgili analizler 203 kişi üzerinden yapılmıştır.

### **Veri Toplama Süreci**

ALES'in 2017 yılı sonbahar döneminde ve 2018 yılı ALES/1-2-3 dönemlerinde yer alan maddelere ait madde parametreleri MTK'ya göre elde edilmiştir. Elde edilen bu parametrelerle CONCERTO programı yardımıyla hazırlanan gerçek zamanlı BBT katılımcılara uygulama linki gönderilerek uygulanmıştır. BBT sonunda elde edilen yetenek değerlerinin ALES puanına yakın dönüşümlerinin yer aldığı tablo da katılımcılara çevrimiçi veri toplama araçları yardımıyla sunulmuştur.

### **Veri Toplama Araçları**

Gerçek zamanlı BBT hazırlamak için ALES'in 2017 yılı sonbahar dönemi ve 2018 yılı ALES/1-2-3 oturumlarında yer alan maddeler kullanılmıştır. Sayısal ve sözel yetenek testi iki ayrı test olarak ele alınmıştır. ALES'in sayısal ve sözel testlerde 50'şer madde bulunmaktadır. Bu maddeler 3 sayısal, 3 sözel alan uzmanı tarafından ölçtüğü becerilere göre sınıflandırılmıştır. Sayısal test için temel işlem, sözel sayı problemi çözme, rutin problem çözme, olasılıksal düşünme, tablo ve grafik okuma, rutin olmayan problem çözme, geometrik akıl yürütme becerisi ve uzamsal beceri olmak üzere 8 başlık altında toplanmıştır. Sözel test için cümlede anlamı tamamlama, cümleden anlam çıkarma, cümle veya metin oluşturma, metindeki anlam akışını sıralama, metindeki anlam bütünlüğünü koruma, okuduğu metni anlama ve yorumlama ve sözel akıl yürütme becerisi olmak üzere 7 başlık altında toplanmıştır.

ALES'in sayısal ve sözel testinde yer alan maddelerin becerilere göre dağılımları tablo 1 ve tablo 2'de yer almaktadır. Maddelerin becerilere göre frekansları ve yüzdeleri hesaplanmıştır. Hazırlanan BBT uygulamasında da bu yüzdeler dikkate alınarak oransal içerik dengeleme yapılmıştır. Her beceriden en az 2 madde olmak üzere sayısal test için 16 madde, sözel test için de 14 madde sorulacak en az madde sayısı olarak belirlenmiştir. Aynı zamanda CONCERTO'nun kullandığı catR paketinde de içerik dengeleme olarak "A-Stratified sampling-tabakalı örnekleme" kullanılmaktadır.

**Tablo 1**

*ALES'in Sayısal Maddelerinin Becerilere Göre Dağılımı*

Beceri Adı	2017	2018-1	2018-2	2018-3	Sınavda
	Sonbahar Madde Sayısı	Madde Sayısı	Madde Sayısı	Madde Sayısı	Sorulma Yüzdesi
Temel İşlem	13	13	14	13	27
Sözel Sayı Problemi	3	2	1	2	4
Rutin Problem Çözme	8	9	8	8	17
Olasılıksal Düşünme	2	2	2	2	4
Tablo ve Grafik Okuma	4	2	6	3	8
Rutin Olmayan Problem Çözme	12	14	11	14	26
Geometrik Akıl Yürütme	4	4	4	4	8
Uzamsal	4	4	4	4	8

Madde havuzunda sayısal ve sözel testler için ALES'in 4 döneminde kullanılmış olan toplam 200'er madde yer almaktadır. Bu maddelerin bazıları ortak metne dayalı fakat bağımsız çözülmesi gereken maddelerdir. Bu sebeple BBT uygulamasında bu maddeler tek tek havuza alınmış ve her seferinde ortak olan metin bireye yeniden verilmiştir. Ortak metne dayalı maddelerin analizleri birbirinden bağımsız yapılmıştır. ALES'in kağıt-kalem uygulamasında sayısal ve sözel test için 50 madde adaylara verildiği için testin bitirme kuralı olarak en fazla



50 maddenin katılımcıya sorulması ya da standart hatanın 0,30'un altına düşmesi belirlenmiştir.

**Tablo 2**

*ALES'in Sözel Maddelerinin Becerilere Göre Dağılımı*

Beceri Adı	2017	2018-1	2018-2	2018-3	Sınavda Sorulma Yüzdesi
	Sonbahar Madde Sayısı	Madde Sayısı	Madde Sayısı	Madde Sayısı	
Cümlede anlamı tamamlama	4	5	5	5	10
Cümleden anlam çıkarma	3	3	3	3	6
Cümle veya metin oluşturma	4	4	4	4	8
Metindeki anlam akışını sıralama	2	3	3	3	6
Metindeki anlam bütünlüğü koruma	3	3	3	3	6
Okuduğu metni anlama ve yorumlama	26	24	24	24	49
Sözel akıl yürütme	8	8	8	8	16

Tablo 3'te bireylerin sayısal testinde yanıtladıkları madde sayılarının frekansları verilmiştir.

**Tablo 3**

*Sayısal Testinde Yanıtlanan Madde Sayılarının Frekans Dağılımları*

Yanıtlanan Madde Sayısı	f	%
39	1	0,42
38	1	0,42
37	2	0,84
36	2	0,84
35	1	0,42
34	2	0,84
32	3	1,26
31	3	1,26

29	4	1,68
28	2	0,84
27	3	1,26
26	4	1,68
25	8	3,36
24	8	3,36
23	8	3,36
22	8	3,36
21	11	4,62
20	24	10,08
19	15	6,30
18	33	13,87
17	36	15,13
16	59	24,79
Toplam	238	100

Sayısal testinde en çok 39 madde en az da 16 madde ile yetenek kestirilmiştir. En az çözülebilecek madde sayısı 16 olduğu için test en çok 16 madde ile sonlanmıştır.

Tablo 4'te bireylerin sözel testinde yanıtladıkları madde sayılarının frekansları verilmiştir.

#### **Tablo 4**

##### *Sözel Testinde Yanıtlanan Madde Sayılarının Frekans Dağılımları*

Yanıtlanan Madde Sayısı	f	%
37	1	0,49
36	1	0,49
34	1	0,49
32	1	0,49
31	1	0,49
28	2	0,99
27	1	0,49
26	2	0,99
25	4	1,97
24	3	1,48

23	5	2,46
22	1	0,49
21	3	1,48
20	8	3,94
19	18	8,87
18	38	18,72
17	35	17,24
16	51	25,12
15	27	13,30
Toplam	203	100

Sözel testinde en çok 37 madde en az da 15 madde ile yetenek kestirilmiştir. Test en çok 16 madde ile sonlanmıştır.

Kağıt-kalem uygulaması ALES'te toplam süre 150 dakikadır. Adaylar bu süreyi sayısal ve sözel testlerde istedikleri gibi kullanabildikleri için sayısal ve sözel testler için ayrı ayrı süre kısıtlamasına gidilmemiştir. Daha önceki gerçek BBT uygulamaları da göz önüne alınınca, testin kağıt-kalem uygulamasından daha kısa sürede bitebileceği düşünüldüğünden katılımcılara süre ile ilgili bir kısıtlama yapılmamıştır. Gerçek ALES uygulamasında adaylara toplam 100 madde sorulmakta ve 150 dakika verilmektedir. Bu sebeple ALES'in hız testi olması yerine bir yetenek testi olmasından söz edilebilir.

Gerçek BBT uygulamasında Bayes Modeli (BM) yetenek kestirim yöntemi olarak kullanılmış ve metrik sabiti olarak  $D=1,702$  alınmıştır. Madde seçme yöntemi olarak da Maksimum Fisher Bilgi (MFI) kullanılmıştır. Testi alan katılımcılara, test sonunda, sayısal ve sözel için ayrı ayrı yetenekleri ve her testte cevapladıkları madde sayıları verilmiştir. Test sonunda adaylara ayrı bir çevrimiçi form sunulmuş ve bu formda BBT sonunda kestirilen yeteneklerinin ALES'in sayısal, sözel ve eşit ağırlık puanlarına dönüşümlerinin yer aldığı tablo verilmiştir. İsteyen katılımcılar bu form aracılığıyla BBT ile elde edilen yeteneklerinin ALES puanı karşılığını elde edebilmişlerdir. Bu dönüşümler ALES'e giren adayların MTK ile elde edilen yetenek dağılımlarının normal dağılıma dönüştürüldükten sonra ÖSYM'nin ALES kılavuzunda yazan puan hesaplama formülüne göre yapılmıştır. Formül aşağıda verilmiştir.

Sayısal puanı için sayısal yetenek 0,75 ve sözel yetenek 0,25; sözel puanı için sayısal yetenek 0,25 ve sözel yetenek 0,75 ve eşit ağırlık puanı için sayısal yetenek 0,50 ve sözel yetenek 0,50 olarak ağırlıklandırılmıştır. Bu dönüşümler katılımcıyı bilgilendirme amacıyla yapılmış ve araştırmada kullanılmamıştır.

$$\text{ALES Puanı} = 70 + \frac{30 [ 2 (AP - X) - S ]}{2 (B - X) - S}$$

ALES : Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı

AP : Adayın ağırlıklı puanı

X : Ağırlıklı puan ortalaması

S : Ağırlıklı puanların standart sapması

B : Ağırlıklı puanların en büyüğü

### **Verilerin Analizi**

BBT uygulamasını hazırlamadan önce, kullanılacak maddelerin MTK varsayımlarını karşılayıp karşılamadığı test edilmiştir. ALES'in sayısal ve sözel yetenek testleri analizlerde iki ayrı test olarak ele alınmıştır.

Araştırmanın alt problemlerine ait bulguların elde edilmesinde Pearson Momentler Çarpımı Korelasyon Katsayısı, bağımsız örneklem t testi, Rasch Model, 2 PLM ve lognormal cevaplama süresi modeli kullanılmıştır. Bu analizler için SPSS 25 ve R Studio programlarından yararlanılmıştır. LNCSM analizleri için LNIRT paketi kullanılmıştır.

### **Tek Boyutluluk**

ALES'te sayısal ve sözel yetenek birer boyut olarak varsayılmış ve bu testlerin tek boyutlu olup olmadığını araştırmak için MPLUS 7 yardımıyla doğrulayıcı faktör analizi yapılmıştır. Kategorik veriler için analiz yöntemi olarak WLSMV (weighted least squares mean and variance adjusted) kullanılmıştır. Bu sebeple model uyum indekslerinden ki-kare, RMSEA (root mean square error of approximation), CFI (comparative fit index) ve TLI (Tucker-Lewis

Index) elde edilmiştir. Veri seti çok büyük olduğu için ki-kare değeri raporlanamamıştır. Diğer model uyum indeksleri tablo 5'te sunulmuştur.

**Tablo 5**

*Doğrulayıcı Faktör Analizine Ait Model Uyum İndeksleri*

Test	RMSEA	CFI	TLI
2017 Sayısal	0,069	0,91	0,90
2017 Sözel	0,069	0,76	0,75
2018-1 Sayısal	0,067	0,89	0,89
2018-1 Sözel	0,072	0,75	0,74
2018-2 Sayısal	0,070	0,92	0,91
2018-2 Sözel	0,073	0,78	0,78
2018-3 Sayısal	0,073	0,87	0,87
2018-3 Sözel	0,075	0,74	0,73

RMSEA değerleri  $<0,08$  olduğu için kabul edilebilir uyum göstermektedir (Jöreskog ve Sörbom, 1993). CFI değeri  $>0,90$  (Çelik ve Yılmaz, 2013) ve TLI değeri de  $>0,90$  (Bentler ve Bonett, 1980) olduğu zaman kabul edilebilir uyuma işaret etmektedir. CFI ve TLI ki-kare değerine bağlı birer fonksiyondur. Bu veri setlerinde de ki-kare değeri oldukça büyük olduğu için istenilen aralıkta yer almamaktadır. Ancak, her test için kurulan modelde maddelerin yükleri 0,30'un üzerinde olduğundan (Comrey ve Lee, 1992) testlerin alt boyutları olsa da toplanabilen tek boyutlu bir üst yapıya sahip olduğu söylenebilir.

**Yerel Bağımsızlık**

Yerel madde bağımsızlığı için kullanılan ölçütlerden birisi de Yen'in Q3 indeksidir (Yen, 1984). Q3, bütün test değerlendirildikten sonra iki madde arasındaki performansı dikkate alan bir korelasyon olduğu için yorumlaması kolaydır. Yerel bağımsız olan maddelerin korelasyonlarının 0 civarında çıkması beklenir ancak ölçme hatasından dolayı 0 ile 0,1

arasında değişebilmektedir. Çalışmalarda  $|Q3|$  değerinin 0,2'den büyük çıkması madde çiftlerinin yerel bağımlı olabileceğini göstermiştir (Ercikan ve ark., 1998).

2017 sonbahar ALES'i için Q3 değeri -0,01 (SRMR=0,03); 2018-1 dönemi ALES'i için Q3 değeri -0,01 (SRMR=0,04); 2018-2 dönemi ALES'i için Q3 değeri -0,01 (SRMR=0,04) ve 2018-3 dönemi ALES'i için Q3 değeri -0,01 (SRMR=0,04) olarak bulunmuştur. Bu durumda kullanılan veri setlerinin yerel bağımsızlık varsayımını karşıladığı sonucuna ulaşılmıştır.

### **Model Veri Uyumu**

BBT uygulaması hazırlamak için kullanılacak madde parametreleri R studio programı yardımıyla elde edilmiştir. Rasch ve 2 PL model kullanılarak madde parametreleri ve yetenek parametreleri kestirilmiştir. Model veri uyumunu incelemek için -2LL, AIC ve BIC değerleri kullanılmıştır. Rasch ve 2 PL model için hesaplanan uyum indekslerinin ikili karşılaştırmalarına ait ANOVA değerleri tablo 6'da verilmiştir.

**Tablo 6**

#### *Rasch Model ve 2 PLM İçin Model Veri Uyumu İncelemesi*

Test	Model	-2LL	AIC	BIC	sd	p
2017	Rasch	-5385351	10770803	10771329	49	0*
Sayısal	2 PLM	-5251960	10504119	10505150		
2017	Rasch	-4774995	9550093	9550618	49	0*
Sözel	2 PLM	-4708516	9417233	9418264		
2018-1	Rasch	-6999331	13998765	13999301	49	0*
Sayısal	2 PLM	-6874228	13748656	13749708		
2018-1	Rasch	-5834301	11668704	11669240	49	0*
Sözel	2 PLM	-5781849	11563897	11564949		
2018-2	Rasch	-3202988	6406079	6406573	49	0*
Sayısal	2 PLM	-3151075	6302349	6303318		
	Rasch	-2578655	5157411	5157906	49	0*

2018-2	2 PLM	-2551803	5103806	5104775		
Sözel						
2018-3	Rasch	-2721675	5443453	5443945	49	0*
Sayısal	2 PLM	-2675958	5352117	5353083		
2018-3	Rasch	-2423311	4846725	4847218	49	0*
Sözel	2 PLM	-2400442	4801083	4802050		

\* $p < 0,05$ ; sd: Serbestlik Derecesi

Tablo 6 incelendiğinde bütün testler için -2LL, AIC ve BIC değerlerinde 2 PLM için azalma vardır. ANOVA sonucunda ise bu farklılığın anlamlı olduğu görülmektedir ( $p < 0,05$ ). Bu sebeple BBT uygulamasında öncelikle 2 PLM ile elde edilen madde kestirimleri kullanılmıştır. Ancak bu maddelerin kullanıldığı BBT'nin pilot uygulamasında özellikle sözel testinde bireylere sıklıkla 50 madde gelmiştir. Yetenek kestiriminde standart hata 0,30'un altına düşmemiştir. Bu sebeple Rasch model ile elde edilen madde parametrelerinin kullanıldığı yeni bir BBT hazırlanmış ve pilot uygulamada sayısal ve sözel testlerde bireylere farklı sayılarda madde geldiği görülmüştür. Araştırmanın uygulamasına Rasch model ile elde edilen madde parametrelerinin kullanıldığı BBT ile devam edilmiştir.

### ***Madde Parametrelerinin Değişmezliği***

Madde parametrelerinin değişmezliğinin araştırılması için iki yol izlenmiştir. Öncelikle ALES'e giren bütün adaylar rastgele iki yarıya bölünmüş ve bu adaylardan Rasch model kullanılarak madde parametreleri elde edilmiştir. Daha sonra bu madde parametreleri arasındaki korelasyon katsayısı Pearson Momentler Çarpımı ile hesaplanmıştır. İkinci olarak ise her dönemin ALES adaylarından rastgele 10 000 kişilik örneklem seçilmiş, bu örneklem de iki yarıya bölünerek 5000'er kişinin madde parametreleri kestirilmiştir. 5000'er kişilik iki gruptan elde edilen madde parametreleri arasındaki korelasyon katsayıları hesaplanmıştır.

**Tablo 7***Madde Parametrelerinin Değişmezliğine Ait Korelasyonlar*

Test	Örneklem	b parametresi
2017	Tüm örneklem	0,99*
Sayısal	10 000 kişi	0,99*
2017	Tüm örneklem	0,99*
Sözel	10 000 kişi	0,99*
2018-1	Tüm örneklem	0,99*
Sayısal	10 000 kişi	0,99*
2018-1	Tüm örneklem	0,99*
Sözel	10 000 kişi	0,99*
2018-2	Tüm örneklem	0,99*
Sayısal	10 000 kişi	0,99*
2018-2	Tüm örneklem	0,99*
Sözel	10 000 kişi	0,99*
2018-3	Tüm örneklem	0,99*
Sayısal	10 000 kişi	0,99*
2018-3	Tüm örneklem	0,99*
Sözel	10 000 kişi	0,99

\* $p < 0,001$

Tablo 7'de araştırmada kullanılan ALES maddeleri için bütün örnekleme ve rastgele seçilen 10 000 kişilik örnekleme tesadüfi olarak ikiye bölerek yapılan korelasyon analizi sonuçları bulunmaktadır. b parametrelerine ait bütün korelasyonlar 0,99 ( $p < 0,001$ ) olarak hesaplanmıştır. Çok yüksek korelasyona işaret eden bu katsayı, rastgele seçilen örneklemlerde de benzer sonuçlar çıktığını göstermektedir. MTK'nın varsayımlarından olan madde parametrelerinin değişmezliği varsayımı da sağlanmış olmaktadır.



### **Yetenek Parametrelerinin Değişmezliği**

Yetenek parametrelerinin değişmezliğini incelemek için öncelikle ALES'e giren adayların sayısal testine verdikleri cevaplardan sayısal yetenekleri daha sonra da sözel testine verdikleri cevaplardan sözel yetenekleri Rasch model ile kestirilmiştir. Elde edilen bu iki yetenek arasındaki ilişki Pearson Momentler Çarpımı ile hesaplanmıştır. Ayrıca ALES'e giren adayların için sayısal ve sözel testler ilk 25 ve son 25 madde olmak üzere ikiye ayrılmış ve bu iki yarıdan da adayların yetenekleri kestirilmiştir. İki yarıdan elde edilen yetenekler arasındaki korelasyon katsayıları hesaplanmıştır.

**Tablo 8**

#### *Yetenek Parametrelerinin Değişmezliğine Ait Korelasyonlar*

Madde Seti	$\theta$	Madde Seti	$\theta$
2017 Sayısal Testi- Sözel Testi	-0,23**	2018-2 Sayısal Testi- Sözel Testi	-0,12**
2017 Sayısal İlk 25 Madde-Son 25 Madde	-0,10**	2018-2 Sayısal İlk 25 Madde-Son 25 Madde	-0,01*
2017 Sözel İlk 25 Madde-Son 25 Madde	-0,32**	2018-2 Sözel İlk 25 Madde-Son 25 Madde	-0,22*
2018-1 Sayısal Testi- Sözel Testi	-0,19**	2018-3 Sayısal Testi- Sözel Testi	-0,21*
2018-1 Sayısal İlk 25 Madde-Son 25 Madde	-0,07**	2018-3 Sayısal İlk 25 Madde-Son 25 Madde	-0,10*
2018-1 Sözel İlk 25 Madde-Son 25 Madde	-0,28**	2018-3 Sözel İlk 25 Madde-Son 25 Madde	-0,29*

\* $p < 0,05$ ; \*\* $p < 0,001$

Tablo 8'de sayısal ve sözel testinden, sayısal testinin ilk 25 maddesi ile son 25 maddesinden ve sözel testinin ilk 25 maddesi ile son 25 maddesinden kestirilen yetenekler arası korelasyonlar yer almaktadır. Korelasyonlar -0,10 ( $p < 0,05$ ) ile -0,32 ( $p < 0,001$ ) arasında

değişmektedir. Çok düşük korelasyona işaret eden bu katsayılar, rastgele seçilen maddelerden kestirilen yeteneklerin farklı olduğunu göstermektedir. ALES'e giren adayların farklı puan türlerinden girdiğini, sayısal ve sözel testinde de farklı becerilerin ölçüldüğü göz önüne alınırsa bunun beklendik bir durum olduğundan söz edilebilir.

### ***Testin Güvenirliđi***

ALES'in kađıt-kalem uygulamasının iç tutarlılık güvenirlilik katsayısı olarak Kuder-Richardson 20 (KR-20) hesaplanmıřtır. ALES 2017 sonbahar dönemi için sayısal testinin 0,96 ve sözel testinin 0,92; 2018-1 dönemi için sayısal testinin 0,95 ve sözel testinin 0,93; 2018-2 dönemi için sayısal testinin 0,96 ve sözel testinin 0,91; 2018-3 dönemi için sayısal testinin 0,95 ve sözel testinin 0,94 olarak elde edilmiřtir.

## Bölüm 4

### Bulgular, Yorumlar ve Tartışma

Bu bölümde araştırmanın problem ve alt problemlerini araştırmak amacıyla yapılan analiz sonuçlarına yer verilmiştir.

#### Araştırmada Elde Edilen Değerlere İlişkin Betimleyici İstatistikler

Araştırma bulgularının daha iyi yorumlanabilmesi için sayısal ve sözel testlerin cevaplandırılmasına ilişkin betimsel analizler verilmiştir. Tablo 9'da sayısal testinde bireylerin aldığı madde sayısına, madde bazında maddeyi cevaplarken harcadıkları ortalama süreye, teste harcadıkları toplam süreye ve kestirilen yetenek parametresine ait betimleyici istatistikler verilmiştir. Betimleyici istatistikler bireylerin aldığı test bazında hesaplanmıştır. Süreler saniye olarak verilmiştir.

**Tablo 9**

*Sayısal Testinde Bireylerin Cevapladıkları Sorulara Ait Betimsel Analiz Sonuçları*

	En düşük değer	En yüksek değer	$\bar{X}$	ss
MS	16	39	20	4,99
b par. ort.	-0,48	2,47	0,97	0,83
MCSO	2,18	558,57	139,27	92,57
TS	39,28	9252	2794,93	1976,59
YD	-1,64	4,00	1,22	1,41

*Not: n = 238. MS: Madde Sayısı; b par. ort.: uygulanan maddelerin b parametrelerinin ortalaması; MCSO: Madde Cevaplama Süresi Ortalaması (sn.); TS: Toplam Süre (sn.); YD: Yetenek Düzeyi*

Tablo 9'a göre bir bireye uygulanan sayısal testteki madde sayısı en az 16 iken en çok 39'dur. Ortalama 20 madde uygulanmıştır (ss= 4,99). Bir testteki madde cevaplama süresi ortalaması 2,18 ile 558,57 saniye arasında değişmektedir. Bir test için bir maddeyi cevaplama süresi ortalaması 139,27 saniyedir (ss=92,57). Madde cevaplama süresinin ortalamasının

çarpıklık katsayısı 0,87 ( $SH=0,16$ ) ve basıklık katsayısı 1,75'tir ( $SH=0,31$ ). Bir test için harcanan toplam süreler 39,28 ile 9252 saniye arasında değişmektedir. Bir test için toplam süre ortalaması 2794,93 saniyedir ( $ss=1976,59$ ). Toplam sürenin çarpıklık katsayısı 0,81 ( $SH=0,16$ ) ve basıklık katsayısı 0,56'dır ( $SH=0,31$ ).

Tablo 10'da sözel testinde bireylerin aldığı madde sayısına, madde bazında maddeyi cevaplarken harcadıkları ortalama süreye, teste harcadıkları toplam süreye ve kestirilen yetenek parametresine ait betimleyici istatistikler verilmiştir.

**Tablo 10**

*Sözel Testinde Bireylerin Cevapladıkları Sorulara Ait Betimsel Analiz Sonuçları*

	En düşük değer	En yüksek değer	$\bar{X}$	ss
MS	15	37	18	3,67
b par. ort.	-1,69	0,71	-0,01	0,56
MCSO	1,94	333,11	93,87	59,56
TS	2432	39552	16724,40	6260,26
YD	-3,15	1,93	-0,04	0,99

*Not: n = 203. MS: Madde Sayısı; b par. ort.: uygulanan maddelerin b parametrelerinin ortalaması; MCSO: Madde Cevaplama Süresi Ortalaması (sn.); TS: Toplam Süre (sn.); YD: Yetenek Düzeyi*

Tablo 10'a göre bir bireye uygulanan sözel testindeki madde sayısı en az 15 iken en çok 37'dir. Ortalama 18 madde uygulanmıştır ( $ss= 3,67$ ). Bir testteki madde cevaplama süresi ortalaması 1,94 ile 333,11 saniye arasında değişmektedir. Bir test için bir maddeyi cevaplama süresi ortalaması 93,87 saniyedir ( $ss=59,56$ ). Madde cevaplama süresinin ortalamasının çarpıklık katsayısı 0,65 ( $SH=0,17$ ) ve basıklık katsayısı 1,25'tir ( $SH=0,34$ ). Bir test için harcanan toplam süreler 2432 ile 39552 saniye arasında değişmektedir. Bir test için toplam süre ortalaması 16724,40 saniyedir ( $ss=6260,26$ ). Toplam sürenin çarpıklık katsayısı 0,01 ( $SH=0,17$ ) ve basıklık katsayısı 0,40'tır ( $SH=0,34$ ).

Araştırmada sayısal testinde bireylere uygulanan maddelerin b parametrelerine ve madde cevaplama süresine ait betimleyici istatistikleri tablo 11'de verilmiştir.

**Tablo 11**

*Sayısal Testinde Maddelere Ait Betimsel Analiz Sonuçları*

	En düşük değer	En yüksek değer	$\bar{X}$	ss
b par.	-2,16	4,48	1,04	1,09
MCS	1,26	2551,03	139,37	174,35

*Not: n = 4767. b par: b parametresi; MCS: Madde Cevaplama Süresi (sn.)*

Tablo 11' göre sayısal testinde uygulanan maddelerin b parametresi -2,16 ile 4,48 arasında değişmektedir. Maddelerin b parametrelerinin ortalaması 1,04'tür ( $ss=1,09$ ). Madde cevaplama süresi 1,26 ile 2551,03 saniye arasında değişmektedir. Madde cevaplama süresinin ortalaması 139,37'dir ( $ss=174,35$ ).

Araştırmada sözel testinde bireylere uygulanan maddelerin b parametrelerine ve madde cevaplama süresine ait betimleyici istatistikler tablo 12'de verilmiştir.

**Tablo 12**

*Sözel Testinde Maddelere Ait Betimsel Analiz Sonuçları*

	En düşük değer	En yüksek değer	$\bar{X}$	ss
b par.	-4,41	2,38	-0,06	0,84
MCS	1,13	2725,90	89,92	124,21

*Not: n = 3679. b par: b parametresi; MCS: Madde Cevaplama Süresi (sn.)*

Tablo 12'ye göre sözel testinde uygulanan maddelerin b parametresi -4,41 ile 2,38 arasında değişmektedir. Maddelerin b parametrelerinin ortalaması -0,06'dır ( $ss=0,84$ ). Madde

cevaplama süresi 1,13 ile 2725,90 saniye arasında değişmektedir. Madde cevaplama süresinin ortalaması 89,92'dir ( $ss=124,21$ ).

Araştırmanın bulgularına alt problemler sırasıyla ele alınarak yer verilmiştir.

**1. Alt Problem “Yetenek düzeyleri ile bireylerin cevapladıkları madde sayıları, madde parametreleri, madde cevaplama süreleri ve test toplam süreleri arasında istatistiksel olarak anlamlı ilişki var mıdır?”**

**a) Sayısal Testi Sonucunda Bireylerden Elde Edilen Değişkenler Arası İlişkiler**

Sayısal testini tamamlayan 238 bireye ait kestirilen yetenek düzeyleri, bireye uygulanan madde sayısı, bir bireyin testindeki b parametrelerinin ortalamaları, madde bazında süre ortalaması ve toplam testte geçen süre arasındaki ilişkiler tablo 13'te verilmiştir.

**Tablo 13**

*Sayısal Testinde Bireylerden Elde Edilen Değişkenler Arasındaki İlişkiler*

Değişkenler	1	2	3	4	5
1. YD	1				
2. MS	0,38*	1			
3. b par. ort.	0,98*	0,33*	1		
4. MCSO	0,57*	-0,01	0,59*	1	
5. TS	0,72*	0,35*	0,70*	0,91*	1

\* $p<0,001$  YD: Yetenek Düzeyi; MS: Madde Sayısı; b par. ort.: uygulanan maddelerin b parametrelerinin ortalaması; MCSO: Madde Cevaplama Süresi Ortalaması (sn.); TS: Toplam Süre (sn.)

Sayısal testini alan bireylerin yetenek düzeyleri ile testteki maddelere ait b parametrelerinin ortalamaları arasında pozitif, çok yüksek ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,98$ ;  $p<0,001$ ); testteki maddelerin süre ortalamaları arasında pozitif, orta düzey ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,57$ ;  $p<0,001$ ) ve testin toplam süresi arasında pozitif, yüksek ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,72$ ;  $p<0,001$ ).

BBT'nin yapısı gereği yetenek arttıkça daha zor madde gelmesinden dolayı yetenek düzeyi ile b parametresi arasında çok yüksek ve pozitif ilişkinin olması beklenen biri durumdur. Yetenek düzeyi arttıkça bireylere daha zor madde geldiği için madde cevaplama süresi ve test için harcanan toplam süre de artmaktadır. Bireyin test için harcadığı toplam süre, bireyin cevaplama gereken madde sayısından da etkilenebilir. Bireylerin cevapladığı madde sayısının yetenek düzeyi ( $r=0,38$ ;  $p<0,001$ ) ve toplam süre ( $r=0,35$ ;  $p<0,001$ ) ile olan ilişkisi düşük ancak b parametresi ile toplam sürenin ilişkisi ( $r=0,70$ ;  $p<0,001$ ) yüksektir. Bu durumda, yetenek düzeyi arttıkça bireylerin test için harcadığı toplam sürenin maddelerin zorluğuna daha bağlı olduğu söylenebilir.

### **b) Sözel Testi Sonucunda Bireylerden Elde Edilen Değişkenler Arası İlişkiler**

Sözel testini tamamlayan 203 bireye ait kestirilen yetenek düzeyleri, bireye uygulanan madde sayısı, bir bireyin testindeki b parametrelerinin ortalamaları, madde bazında süre ortalaması ve toplam testte geçen süre arasındaki ilişkiler tablo 14'te verilmiştir.

**Tablo 14**

#### *Sözel Testinde Bireylerden Elde Edilen Değişkenler Arasındaki İlişkiler*

Değişkenler	1	2	3	4	5
1. YD	1				
2. MS	-0,42**	1			
3. b par. ort.	0,95**	-0,44**	1		
4. MCSO	0,58**	-0,39**	0,57**	1	
5. TS	-0,14*	0,32**	-0,13	-0,13	1

\* $p<0,05$ ; \*\* $p<0,001$  YD: Yetenek Düzeyi; MS: Madde Sayısı; b par. ort.: uygulanan maddelerin b parametrelerinin ortalaması; MCSO: Madde Cevaplama Süresi Ortalaması (sn.); TS: Toplam Süre (sn.)

Sözel testini alan bireylerin yetenek düzeyleri ile testteki maddelere ait b parametrelerinin ortalamaları arasında pozitif, çok yüksek ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,95$ ;  $p<0,001$ ); testteki maddelerin süre ortalamaları arasında pozitif, orta düzey ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,58$ ;  $p<0,001$ ). Bu ilişkiler sayısal testi ile paralellik göstermektedir. Ancak yetenek düzey ve testin toplam süresi arasında negatif, çok düşük ve

istatistiksel olarak anlamlı olan ilişki ( $r=-014$ ;  $p<0,05$ ) sayısal testine göre ayrılmaktadır. Ayrıca toplam süre ile sadece madde sayısı arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,32$ ;  $p<0,001$ ).

Sözel testinde yer alan maddelerin b parametreleri incelendiğinde, sözel testinin sayısal testine göre daha kolay maddelerden oluştuğu ve madde cevaplama süresi ortalamasının daha düşük olduğu görülmektedir. Bu durum teste harcanan toplam süre ile b parametresinin ortalaması ve madde cevaplama süresinin ortalamasının arasında istatistiksel olarak ilişki olmamasına ( $p>0,05$ ) yansımış olabilir.

**2. Alt Problem “Düşük ve yüksek yetenek düzeylerine göre bireylerin cevapladıkları madde sayıları, madde parametreleri, madde cevaplama süreleri ve test toplam süreleri arasında istatistiksel olarak anlamlı fark var mıdır?”**

**a) Sayısal Testi Sonucunda Yetenek Düzeylerine Göre Değişkenlerin Karşılaştırılması**

Sayısal testini alan 238 birey yetenek düzeylerine göre %27'lik alt – üst gruba ayrılmıştır. Düşük ( $\theta_{ort}=-0,69$ ;  $n=64$ ) ve yüksek ( $\theta_{ort}=2,84$ ;  $n=64$ ) yetenek grubundaki bireylerden elde edilen madde sayısına, b parametreleri ortalamasına, madde cevaplama süresi ortalamasına ve toplam süreye ait karşılaştırmaların yer aldığı bağımsız örneklem t testi sonuçları tablo 15'te verilmiştir.



**Tablo 15**

*Sayısal Testinde Yetenek Düzeyine Göre Değişkenlere Ait Bağımsız Örneklem t Testi*

*Tablosu*

Değişkenler	Gruplar	n	$\bar{X}$	ss	t	df	p	$\eta^2$
MS	Düşük	64	19,47	4,62	-5,67	126	0,00*	0,20
	Yüksek	64	24,69	5,73				
b par ort.	Düşük	64	-0,16	0,24	-49,71	126	0,00*	0,95
	Yüksek	64	1,90	0,22				
MCSO	Düşük	64	51,24	54,77	-12,36	126	0,00*	0,55
	Yüksek	64	184,95	67,02				
TS	Düşük	64	900,78	909,99	-14,83	95,87	0,00*	0,64
	Yüksek	64	4500,34	1715,11				

\* $p < 0,001$  YD: Yetenek Düzeyi; MS: Madde Sayısı; b par. ort.: uygulanan maddelerin b parametrelerinin ortalaması; MCSO: Madde Cevaplama Süresi Ortalaması (sn.); TS: Toplam Süre (sn.)

Tablo 15'te yer alan sonuçlara göre sayısal testini alan bireylerin yetenek düzeyine göre oluşturulan gruplarda bireylerin testte aldıkları madde sayısı anlamlı farklılık göstermektedir ( $t(126)=-5,67$ ;  $p < 0,001$ ) ve  $\eta^2$  değeri 0,20 olarak hesaplanmıştır. Bireylerin testte aldıkları madde sayısının yüksek düzeyde öneme sahip olduğu söylenebilir. Sayısal testini alan bireylerin yetenek düzeyine göre oluşturulan gruplarda bireylerin testte aldıkları maddelerin b parametrelerinin ortalaması anlamlı farklılık göstermektedir ( $t(126)=-49,71$ ;  $p < 0,001$ ) ve  $\eta^2$  değeri 0,95 olarak hesaplanmıştır. Bireylerin testte aldıkları maddelerin b parametrelerinin ortalamasının yüksek düzeyde öneme sahip olduğu söylenebilir. Sayısal testini alan bireylerin yetenek düzeyine göre oluşturulan gruplarda bireylerin testte aldıkları maddelere harcadıkları sürelerin ortalaması anlamlı farklılık göstermektedir ( $t(126)=-12,36$ ;  $p < 0,001$ ) ve  $\eta^2$  değeri 0,55 olarak hesaplanmıştır. Bireylerin testte aldıkları maddelere harcadıkları sürelerinin ortalamasının yüksek düzeyde öneme sahip olduğu söylenebilir. Sayısal testini alan bireylerin yetenek düzeyine göre oluşturulan gruplarda bireylerin teste harcadıkları toplam süre anlamlı farklılık göstermektedir ( $t(95,87)=-14,83$ ;  $p < 0,001$ ) ve  $\eta^2$  değeri 0,64 olarak hesaplanmıştır. Bireylerin teste harcadıkları toplam süre yüksek düzeyde öneme sahip olduğu söylenebilir.

BBT'de yüksek yetenek grubunda yer alan bireylere zor maddeler geldiği için b parametrelerinin ortalamaları arasında anlamlı farklılık çıkması beklenen bir durumdur. Madde sayısı, madde cevaplama süresi ortalaması ve teste harcanan toplam süre incelendiği zaman yüksek yetenek grubunda yer alan bireylere ait ortalamaların daha yüksek olduğu görülmektedir. Madde sayısının artmasının ve daha zor maddelerin gelmesinin hem madde cevaplama süresi ortalamasının artmasına hem de teste harcanan toplam sürenin artmasına sebep olduğu söylenebilir.

### **b) Sözel Testi Sonucunda Yetenek Düzeylerine Göre Değişkenlerin Karşılaştırılması**

Sözel testini alan 203 birey yetenek düzeylerine göre %27'lik alt – üst gruba ayrılmıştır. Düşük ( $\theta_{ort}=-1,34$ ;  $n=55$ ) ve yüksek ( $\theta_{ort}=1,93$ ;  $n=56$ ) yetenek grubundaki bireylerden elde edilen madde sayısına, b parametreleri ortalamasına, madde cevaplama süresi ortalamasına ve toplam süreye ait karşılaştırmaların yer aldığı bağımsız örneklem t testi sonuçları tablo 16'da verilmiştir.

**Tablo 16**

*Sözel Testinde Yetenek Düzeyine Göre Değişkenlere Ait Bağımsız Örneklem t Testi Tablosu*

Değişkenler	Gruplar	n	$\bar{X}$	ss	t	df	p	$\eta^2$
MS	Düşük	55	20,98	4,55	2,93	105,52	0,00*	0,07
	Yüksek	56	18,73	3,44				
b par ort.	Düşük	55	-0,73	0,44	-19,75	67,91	0,00**	0,79
	Yüksek	56	0,52	0,16				
MCSO	Düşük	55	42,45	47,24	-8,20	109	0,00**	0,38
	Yüksek	56	121,49	54,08				
TS	Düşük	55	18914,96	6463,52	1,98	109	0,05	0,04
	Yüksek	56	16433,89	6739,23				

\* $p<0,01$ ; \*\* $p<0,001$ ; YD: Yetenek Düzeyi; MS: Madde Sayısı; b par. ort.: uygulanan maddelerin b parametrelerinin ortalaması; MCSO: Madde Cevaplama Süresi Ortalaması (sn.); TS: Toplam Süre (sn.)

Tablo 16'da yer alan sonuçlara göre sözel testini alan bireylerin yetenek düzeyine göre oluşturulan gruplarda bireylerin testte aldıkları madde sayısı anlamlı farklılık göstermektedir

$(t(105,52)=-2,93; p<0,01)$  ve  $\eta^2$  değeri 0,07 olarak hesaplanmıştır. Bireylerin testte aldıkları madde sayısının orta düzeyde öneme sahip olduğu söylenebilir. Sözel testini alan bireylerin yetenek düzeyine göre oluşturulan gruplarda bireylerin testte aldıkları maddelerin b parametrelerinin ortalaması anlamlı farklılık göstermektedir  $(t(67,91)=-19,75; p<0,001)$  ve  $\eta^2$  değeri 0,79 olarak hesaplanmıştır. Bireylerin testte aldıkları maddelerin b parametrelerinin ortalamasının yüksek düzeyde öneme sahip olduğu söylenebilir. Sözel testini alan bireylerin yetenek düzeyine göre oluşturulan gruplarda bireylerin testte aldıkları maddelere harcadıkları sürelerin ortalaması anlamlı farklılık göstermektedir  $(t(109)=-8,20; p<0,001)$  ve  $\eta^2$  değeri 0,38 olarak hesaplanmıştır. Bireylerin testte aldıkları maddelere harcadıkları sürelerinin ortalamasının yüksek düzeyde öneme sahip olduğu söylenebilir. Sözel testini alan bireylerin yetenek düzeyine göre oluşturulan gruplarda bireylerin teste harcadıkları toplam süre anlamlı farklılık göstermemektedir  $(t(109)=1,98; p>0,05)$  ve  $\eta^2$  değeri 0,04 olarak hesaplanmıştır. Bireylerin teste harcadıkları toplam süre düşük düzeyde öneme sahip olduğu söylenebilir.

BBT'de sözel testinde yüksek yetenek grubundaki bireylere, düşük yetenek grubundakilerden anlamlı derecede az madde sorulurken, daha zor madde geldiği için madde cevaplama süresi ortalaması daha yüksektir. Düşük yetenek düzeyindekilere ise çok madde fakat daha az cevaplama süresi gerektiren maddeler sorulduğu için testte harcanan toplam süre sözel testi için yetenek grubuna göre farklılaşmamıştır.

### **3. Alt Problem “Maddenin doğru veya yanlış cevaplanmasına göre madde cevaplama süresinde istatistiksel olarak anlamlı fark var mıdır?”**

Sayısal ve sözel testinde bireylere uygulanan maddelerin b parametreleri ile madde cevaplama süreleri arasındaki ilişkiler ve bu maddeleri doğru veya yanlış yapma durumuna göre ilişkiler incelenmiştir.

#### **a) Sayısal Testi Sonucunda Maddelerden Elde Edilen Değişkenler Arası İlişkiler**

Sayısal testinde uygulanan 4767 maddenin b parametreleri ile maddeleri cevaplama süresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır  $(r=0,35; p<0,001)$ .

Sayısal testinde 2950 madde doğru ve 1817 madde yanlış cevaplanmıştır. Doğru cevaplanan maddelerin b parametreleri ile maddeleri cevaplama süresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,31$ ;  $p<0,001$ ). Yanlış cevaplanan maddelerin b parametreleri ile maddeleri cevaplama süresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,38$ ;  $p<0,001$ ).

b parametresi artınca madde zorluk düzeyi de arttığı için madde cevaplama süresi de artmaktadır.

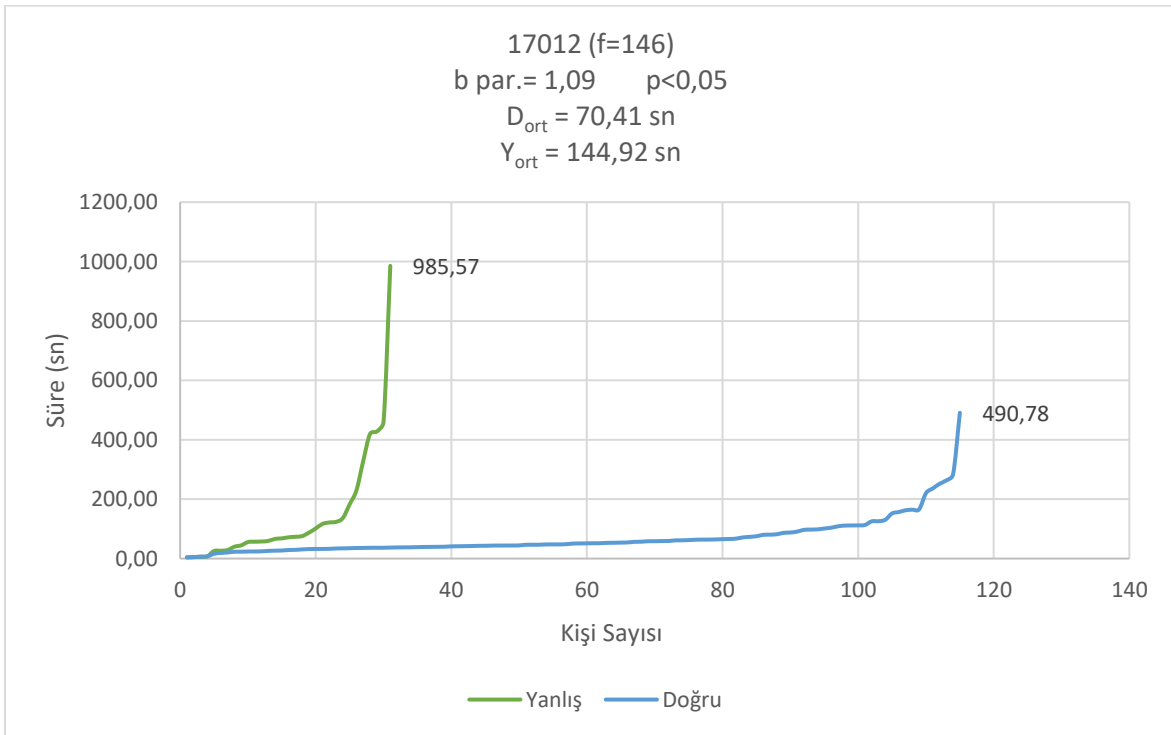
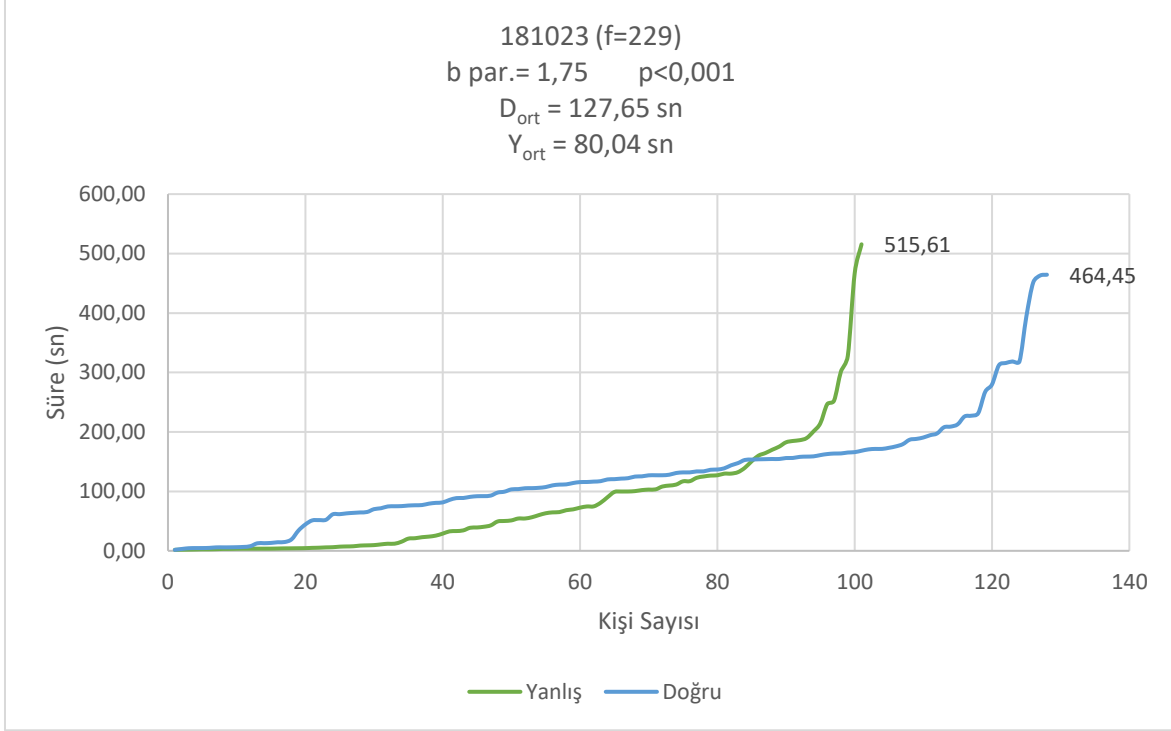
### ***b) Sayısal Testi Sonucunda Madde Cevaplama Süresinin Maddeyi Doğru veya Yanlış Yapma Durumuna Göre Karşılaştırılması***

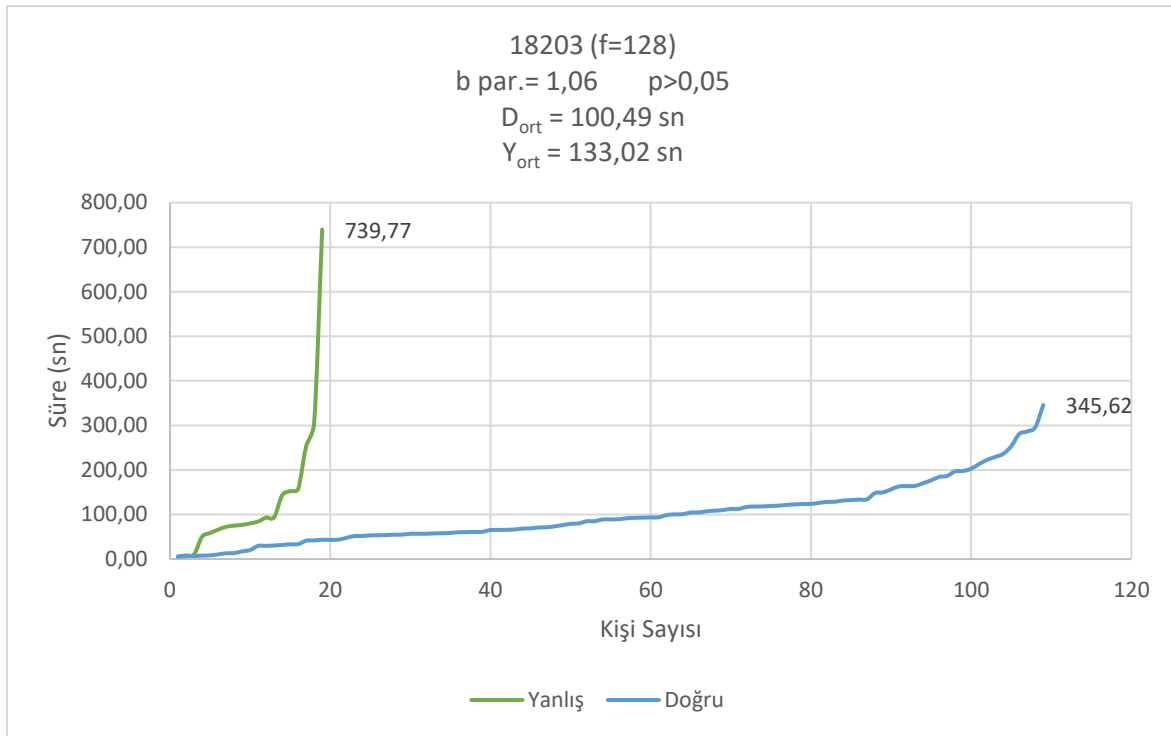
Sayısal testinde en çok tekrar eden 10 tane madde belirlenmiş ve bu maddeleri cevaplarken harcanan süreler maddenin doğru veya yanlış cevaplanma durumuna göre karşılaştırılmıştır. 10 maddenin 7 tanesinde doğru veya yanlış cevaplama durumuna göre istatistiksel olarak anlamlı farklılık varken ( $\bar{X}_{b\ par.}=0,53$ ;  $p<0,05$ ), 3 tanesinde anlamlı farklılık bulunmamaktadır ( $\bar{X}_{b\ par.}=1,62$ ;  $p>0,05$ ). Anlamlı farklılık bulunan 7 maddenin 5 tanesinde ( $\bar{X}_{b\ par.}=0,49$ ) doğru cevaplayanların maddeye ayırdıkları süre yanlış cevaplayanların maddeye ayırdıkları süreden daha fazla iken, 2 tanesinde ( $\bar{X}_{b\ par.}=0,63$ ) durum tam tersidir. Şekil 4'te sayısal testinde en çok tekrar eden 3 madde için madde cevaplama süresinin maddenin doğru veya yanlış cevaplanma durumuna göre değişim grafikleri ve yapılan bağımsız örneklem t testi sonuçları maddelerin b parametreleri ile birlikte verilmiştir. Analizleri yapılan 7 maddenin grafiklerine Ek-A'da yer verilmiştir.

Madde zorluk düzeyi arttıkça, maddede geçirilen süre arttığı için bireylerin fazla zaman geçirdikleri düşüncesi ile maddeyi yanlış yapma eğiliminde de olabileceği ilgili alan yazında yer almaktadır. Yanlış cevaplanan maddelerin b parametrelerinin ortalamaları incelendiğinde, doğru cevaplanan maddelere göre yüksek olduğu görülmektedir. Daha zor olan bu maddelerde adaylar fazla zaman harcadıkları için doğru cevaba ulaşamamaları bile maddeyi boş bırakmamak için işaretlemiş olabilirler.

#### Şekil 4

*Sayısal Testinde Madde Cevaplama Süresinin Maddenin Doğru veya Yanlış Cevaplanma Durumuna Göre Değişim Grafikleri*





### c) Sözel Testi Sonucunda Maddelerden Elde Edilen Değişkenler Arası İlişkiler

Sözel testinde uygulanan 3679 maddenin b parametreleri ile maddeleri cevaplama süresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,32$ ;  $p<0,001$ ).

Sözel testinde 1771 madde doğru ve 1908 madde yanlış cevaplanmıştır. Doğru cevaplanan maddelerin b parametreleri ile maddeleri cevaplama süresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,34$ ;  $p<0,001$ ). Yanlış cevaplanan maddelerin b parametreleri ile maddeleri cevaplama süresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,27$ ;  $p<0,001$ ).

Sayısal testi ile sözel testi benzerlik göstermektedir ve b parametresi artınca madde zorluk düzeyi de arttığı için madde cevaplama süresi de artmaktadır.

### d) Sözel Testi Sonucunda Madde Cevaplama Süresinin Maddeyi Doğru veya Yanlış Yapma Durumuna Göre Karşılaştırılması

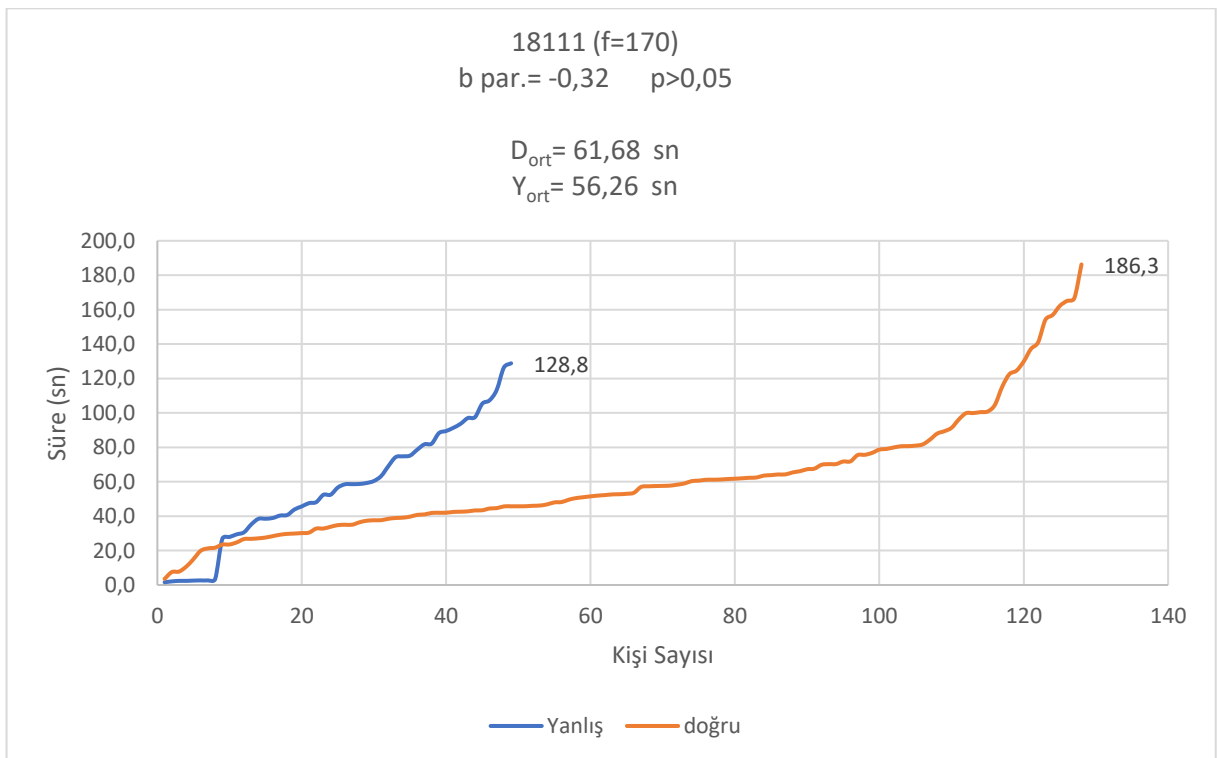
Sözel testinde en çok tekrar eden 10 tane madde belirlenmiş ve bu maddeleri cevaplarken harcanan süreler maddenin doğru veya yanlış cevaplanma durumuna göre karşılaştırılmıştır. 10 maddenin 3 tanesinde doğru veya yanlış cevaplama durumuna göre

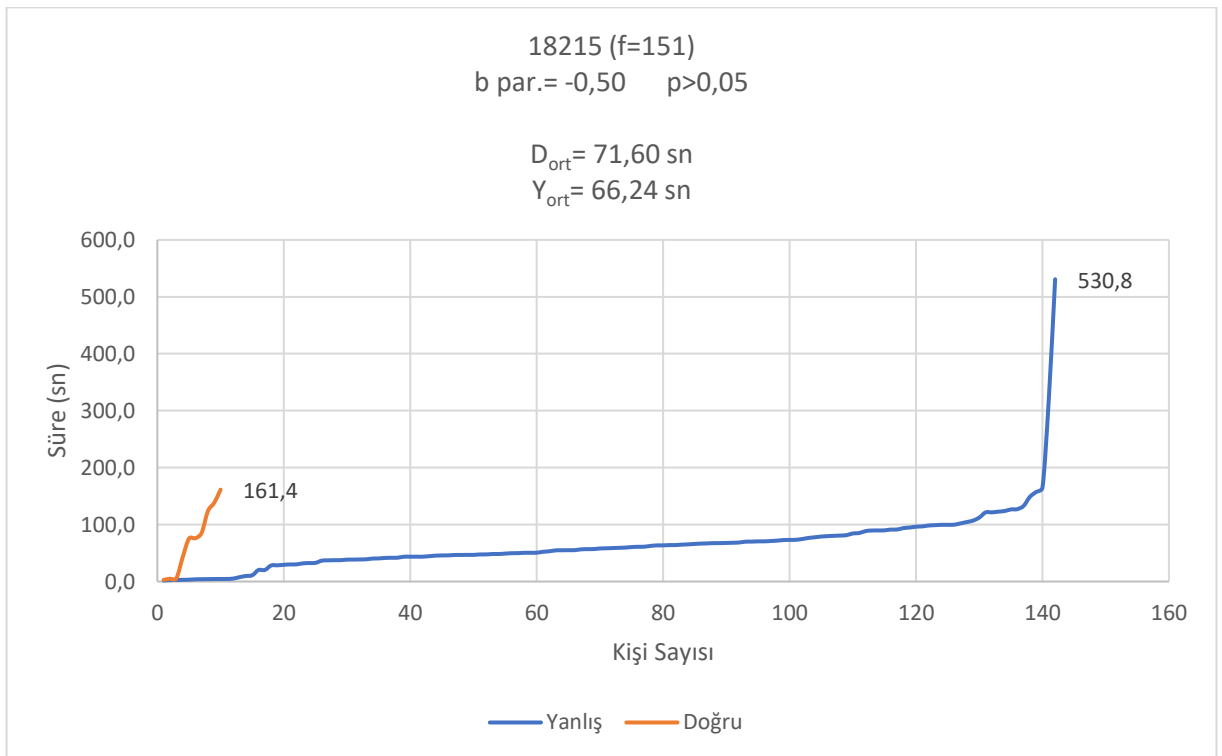
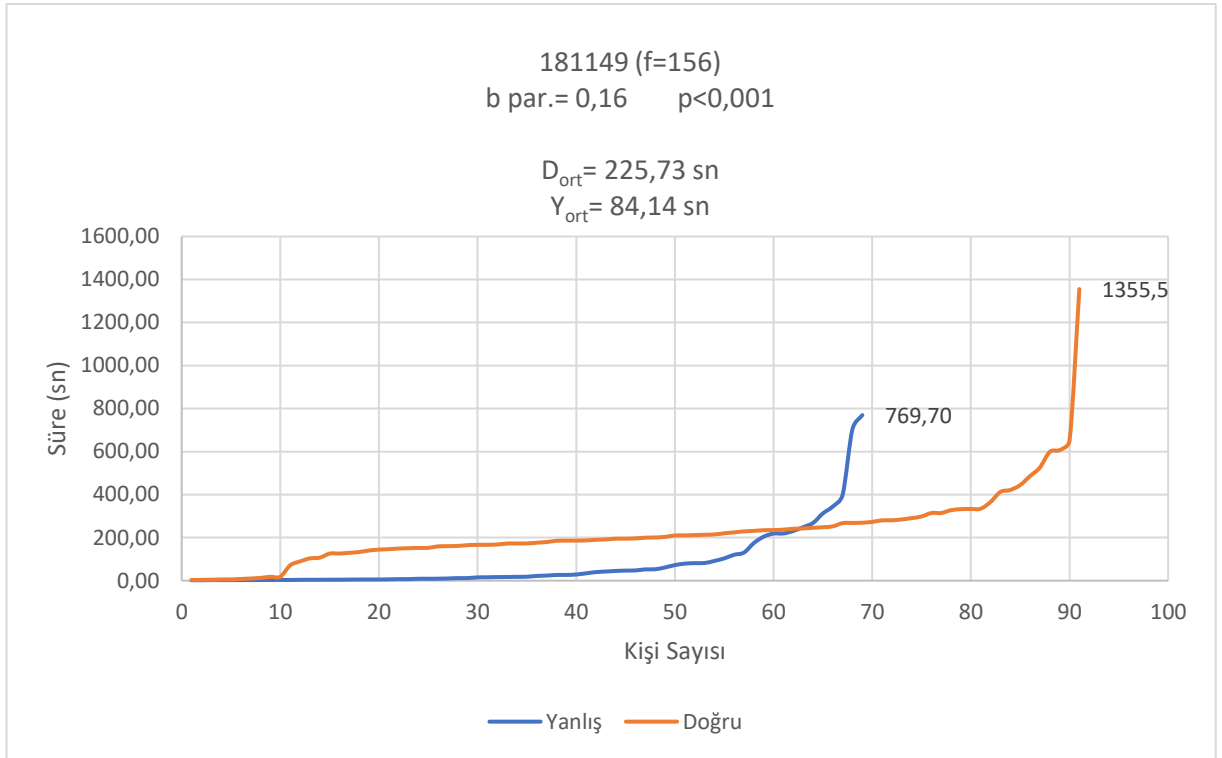
istatistiksel olarak anlamlı farklılık varken ( $\bar{X}_{b\ par.}=0,35; p<0,05$ ), 7 tanesinde anlamlı farklılık bulunmamaktadır ( $\bar{X}_{b\ par.}=0,14; p>0,05$ ). Anlamlı farklılık bulunan 3 maddenin tamamında doğru cevaplayanların maddeye ayırdıkları süre yanlış cevaplayanların maddeye ayırdıkları süreden daha fazladır. Şekil 5'te sözel testinde en çok tekrar eden 3 madde için madde cevaplama süresinin maddenin doğru veya yanlış cevaplanma durumuna göre değişim grafikleri ve yapılan bağımsız örneklem t testi sonuçları maddelerin b parametreleri ile birlikte verilmiştir. Analizleri yapılan 7 maddenin grafiklerine Ek-B'de yer verilmiştir.

Sözel testinde incelenen en çok tekrar eden maddeler orta düzey zorlukta olduğu için doğru veya yanlış cevaplama durumları, maddeyi cevaplamak için harcanan süre için farklılık göstermemektedir.

### Şekil 5

*Sözel Testinde Madde Cevaplama Süresinin Maddenin Doğru veya Yanlış Cevaplanma Durumuna Göre Değişim Grafikleri*







**4. Alt Problem “Madde cevaplama süresini dikkate alan lognormal cevaplama süresi modeli (LNCSM) ile kestirilen madde parametreleri ve madde cevaplama süresini dikkate almayan MTK modelleri ile kestirilen madde parametreleri arasında istatistiksel olarak anlamlı ilişki var mıdır?”**

BBT uygulamasında kullanılan maddelerin Rasch Modele göre elde edilen b parametresi, 2 PL modele göre elde edilen a ve b parametreleri, BBT'nin sayısal ve sözel testinde yer alan maddeler için LNCSM ile kestirilen madde ayırt edicilik, madde güçlük, süre ayırt edicilik ve süre yoğunluk parametreleri arasındaki ilişkiler incelenmiştir. BBT'nin sayısal ve sözel testinde yer alan maddeler için LNCSM ile kestirilen madde ayırt edicilik, madde güçlük, süre ayırt edicilik ve süre yoğunluk parametrelerine ait betimleyici istatistikler Ek-C'de verilmiştir.

***a) Sayısal Testi Sonucunda LNCSM'den ve Cevaplama Süresini Dikkate Almayan MTK Modellerinden Elde Edilen Madde Parametreleri Arasındaki İlişkiler***

BBT'nin sayısal testinde kullanılan 192 maddeye ait LNCSM sonucu elde edilen madde ayırt edicilik, madde güçlük, süre ayırt edicilik ve süre yoğunluk parametreleri ile bu maddelerden MTK'da Rasch modeli sonucu elde edilen b parametresi ve 2 PL model ile elde edilen a ve b parametreleri arasındaki ilişkiler tablo 17'de verilmiştir.

Sayısal testindeki maddelerden LNCSM sonrası elde edilen madde ayırt edicilik parametresi ve 2 PL modelden elde edilen a parametresi arasında istatistiksel olarak anlamlı ilişki yoktur ( $r=0,11$ ;  $p>0,05$ ); süre ayırt edicilik parametresi arasında pozitif, orta düzey ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,47$ ;  $p<0,001$ ); süre yoğunluk parametresi arasında pozitif, çok düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,25$ ;  $p<0,001$ );

LNCSM sonrası elde edilen madde güçlük parametresi ile süre ayırt edicilik parametresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,38$ ;  $p<0,001$ ); süre yoğunluk parametresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,38$ ;  $p<0,001$ ); Rasch modelden elde edilen b parametresi arasında pozitif, çok düşük ve

istatistiksel olarak anlamlı ilişki vardır ( $r=0,25$ ;  $p<0,01$ ) ve 2 PL modelden elde edilen b parametresi arasında pozitif, çok düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,21$ ;  $p<0,01$ ).

**Tablo 17**

*Sayısal Testinde LNCSM ve MTK Modellerinden Elde Edilen Madde Parametreleri Arasındaki İlişkiler*

Değişkenler	1	2	3	4	5	6	7
1. M. Ayırt Ediciliği	1						
2. Madde Güçlüğü	-0,09	1					
3. S. Ayırt Ediciliği	0,47**	0,38**	1				
4. Süre Yoğunluğu	0,25**	0,38**	0,45**	1			
5. b par. (Rasch)	0,28**	0,25*	0,20*	0,42**	1		
6. a par. (2PLM)	0,11	-0,09	0,21*	-0,11	-0,28**	1	
7. b par. (2PLM)	0,32**	0,21*	0,21*	0,42**	0,99**	-0,21*	1

\* $p<0,01$ ; \*\* $p<0,001$

LNCSM sonrası elde edilen süre ayırt edicilik parametresi ile süre yoğunluk parametresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,45$ ;  $p<0,001$ ); Rasch modelden elde edilen b parametresi arasında pozitif, çok düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,20$ ;  $p<0,01$ ) ve 2 PL modelden elde edilen b parametresi arasında pozitif, çok düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,21$ ;  $p<0,01$ ).

LNCSM sonrası elde edilen süre yoğunluk parametresi ile Rasch modelden elde edilen b parametresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,42$ ;  $p<0,001$ ) ve 2 PL modelden elde edilen b parametresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,42$ ;  $p<0,001$ ).

Süre ayırt edicilik parametresi çalışma hızının madde cevaplama olan madde etkisini gösterdiği için madde güçlük ve madde ayırt edicilik parametreleri ile ilişki bulunmuştur. Süre

yoğunluğu parametresi, maddeyi cevaplamak için harcanan ortalama süreyi temsil etmektedir. Bu sebeple madde güçlük parametresi ile ilişkisi bulunmaktadır.

**b) Sözel Testi Sonucunda LNCSM'den ve Cevaplama Süresini Dikkate Almayan MTK Modellerinden Elde Edilen Madde Parametreleri Arasındaki İlişkiler**

BBT'nin sözel testinde kullanılan 192 maddeye ait LNCSM sonucu elde edilen madde ayırt edicilik, madde güçlük, süre ayırt edicilik ve süre yoğunluk parametreleri ile bu maddelerden MTK'da Rasch modeli sonucu elde edilen b parametresi ve 2 PL model ile elde edilen a ve b parametreleri arasındaki ilişkiler tablo 18'de verilmiştir.

**Tablo 18**

*Sözel Testinde LNCSM ve MTK Modellerinden Elde Edilen Madde Parametreleri Arasındaki İlişkiler*

Değişkenler	1	2	3	4	5	6	7
1. M. Ayırt Ediciliği	1						
2. Madde Güçlüğü	-0,19**	1					
3. S. Ayırt Ediciliği	0,52***	-0,03	1				
4. Süre Yoğunluğu	0,63***	-0,10	-0,10	1			
5. b par. (Rasch)	0,09	0,09	-0,20**	0,35***	1		
6. a par. (2PLM)	0,18*	-0,15*	0,35***	-0,05	-0,14	1	
7. b par. (2PLM)	0,10	0,08	-0,20**	0,35***	0,99**	-0,24**	1

\* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$

Sözel testindeki maddelerden LNCSM sonrası elde edilen madde ayırt edicilik parametresi ve 2 PL modelden elde edilen a parametresi arasında pozitif, çok düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,18$ ;  $p < 0,05$ ).

LNCSM sonrası elde edilen madde güçlük parametresi ile süre ayırt edicilik parametresi, süre yoğunluk parametresi, Rasch modelden elde edilen b parametresi ve 2 PL modelden elde edilen b parametresi arasında istatistiksel olarak anlamlı ilişki yoktur ( $p > 0,05$ ).

LNCSM sonrası elde edilen süre ayırt edicilik parametresi ile süre yoğunluk parametresi arasında istatistiksel olarak anlamlı ilişki yoktur ( $r=-0,10$ ;  $p>0,05$ ); Rasch modelden elde edilen b parametresi arasında negatif, çok düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=-0,20$ ;  $p<0,01$ ) ve 2 PL modelden elde edilen b parametresi arasında negatif, çok düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=-0,20$ ;  $p<0,01$ ).

LNCSM sonrası elde edilen süre yoğunluk parametresi ile Rasch modelden elde edilen b parametresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,35$ ;  $p<0,001$ ) ve 2 PL modelden elde edilen b parametresi arasında pozitif, düşük ve istatistiksel olarak anlamlı ilişki vardır ( $r=0,35$ ;  $p<0,001$ ).

Sayısal ve sözel testlerde, 2 PLM ile elde edilen a parametresinin diğer parametreler ile ilişkisinin anlamlı olmaması veya çok düşük olması hazırlanan BBT uygulamasındaki parametrelerin Rasch model ile elde edilmesinden kaynaklanmış olabilir. LNCSM de kullanılan maddelerin parametreleri BBT ile elde edilmiş olduğu için a parametresinin, bu maddeler üzerine etkisi bulunmamaktadır.

Sayısal ve sözel testler için LNCSM ve madde cevaplama süresini dikkate almayan MTK modellerinden elde edilen parametreler arası ilişkiler düşüktür.

##### **5. Alt Problem “Madde cevaplama süresini dikkate alan lognormal cevaplama süresi modeli (LNCSM) ile kestirilen yetenek düzeyleri ve madde cevaplama süresini dikkate almayan MTK modelleri ile kestirilen yetenek düzeyleri arasında istatistiksel olarak anlamlı ilişki var mıdır?”**

BBT uygulaması sonucu bireylerin maddeleri cevaplarken harcadıkları süreler kaydedilmiştir. Bu sürelerin de modele dahil olduğu Lognormal Cevaplama Süresi Modeli (LNCSM) ile bireylerin yetenekleri yeniden kestirilmiş ve BBT uygulaması ile elde edilen yetenekler arasındaki ilişkiler incelenmiştir. BBT'nin sayısal ve sözel testinde yer alan maddeler için LNCSM ile kestirilen yetenek ve hız değerlerine ait betimleyici istatistikler Ek-D'de verilmiştir.

**a) Sayısal Testi Sonucunda LNCSM ve BBT'den Elde Edilen Yetenek Değerleri Arasındaki İlişkiler**

BBT'de sayısal testini tamamlayan 238 bireye ait kestirilen yetenek düzeyleri ve bu bireylerin LNCSM'ye göre kestirilen yetenek ve kişisel hız değerleri arasındaki ilişkiler tablo 19'da verilmiştir.

**Tablo 19**

*Sayısal Testinde LNCSM ve BBT'den Elde Edilen Yetenek ve Hız Değerleri Arasındaki İlişkiler*

Değişkenler	1	2	3
1. LNCSM_yetenek	1		
2. LNCSM_hız	-0,83*	1	
3. BBT_yetenek	0,95*	-0,70*	1

\* $p < 0,001$

Sayısal testinde bireylerin LNCSM sonrası elde edilen yetenek düzeyleri ile LNCSM'den elde edilen kişisel hız değerleri arasında negatif, yüksek ve istatistiksel olarak anlamlı ilişki vardır ( $r = -0,83$ ;  $p < 0,001$ ) ve BBT'den elde edilen yetenek değerleri arasında pozitif, çok yüksek ve istatistiksel olarak anlamlı ilişki vardır ( $r = 0,95$ ;  $p < 0,001$ ). Sayısal testinde bireylerin LNCSM sonrası elde edilen kişisel hız değerleri ile BBT'den elde edilen yetenek değerleri arasında negatif, yüksek ve istatistiksel olarak anlamlı ilişki vardır ( $r = -0,70$ ;  $p < 0,001$ ).

**b) Sözel Testi Sonucunda LGIRT ve BBT'den Elde Edilen Yetenek Değerleri Arasındaki İlişkiler**

BBT'de sözel testini tamamlayan 203 bireye ait kestirilen yetenek düzeyleri ve bu bireylerin LNCSM'ye göre kestirilen yetenek ve kişisel hız değerleri arasındaki ilişkiler tablo 20'de verilmiştir.

**Tablo 20**

*Sözel Testinde LNCSM ve BBT'den Elde Edilen Yetenek ve Hız Değerleri Arasındaki İlişkiler*

Değişkenler	1	2	3
1. LNCSM_yetenek	1		
2. LNCSM_hız	-0,90*	1	
3. BBT_yetenek	0,92*	-0,76*	1

\* $p < 0,001$

Sözel testinde bireylerin LNCSM sonrası elde edilen yetenek düzeyleri ile LNCSM'den elde edilen kişisel hız değerleri arasında negatif, çok yüksek ve istatistiksel olarak anlamlı ilişki vardır ( $r = -0,90$ ;  $p < 0,001$ ) ve BBT'den elde edilen yetenek değerleri arasında pozitif, çok yüksek ve istatistiksel olarak anlamlı ilişki vardır ( $r = 0,92$ ;  $p < 0,001$ ). Sözel testinde bireylerin LNCSM sonrası elde edilen kişisel hız değerleri ile BBT'den elde edilen yetenek değerleri arasında negatif, yüksek ve istatistiksel olarak anlamlı ilişki vardır ( $r = -0,76$ ;  $p < 0,001$ ).

LNCSM ve BBT sonucunda elde edilen yetenek parametreleri arasındaki ilişki çok yüksektir. LNCSM sonucu elde edilen hız parametresi ile LNCSM sonucu elde edilen yetenek parametresinin negatif ilişkili olması beklenmektedir. Hız düştükçe LNCSM modelde yeteneğe karşılık gelen doğruluk parametresi artacağı için BBT ile elde edilen yetenek parametresi ile hız parametresi arasındaki negatif ilişki de beklenen bir durumdur.

## Bölüm 5

### Sonuç ve Öneriler

#### Sonuçlar

Araştırmanın bulgularından elde edilen sonuçlar aşağıda listelenmiştir.

- Sayısal testinde yer alan maddeler daha zordur. Ayrıca, bireylerin yeteneklerinin kestirilmesi için sayısal testinde cevaplamaları gereken madde sayısı daha fazla olduğu için madde süre ortalaması ve testin toplam süresi daha fazladır.
- Sayısal testinde yetenek düzeyi yüksek olanların teste harcadıkları süre daha fazla iken sözel testinde yetenek düzeyi ile teste harcanan süre ilişkili değildir. Yetenek düzeyi ile teste harcanan sürenin ilişkili olma durumu maddelerin zorluğuna bağlı olabilir.
- Sayısal testinde yetenek düzeyi yüksek olanların test boyunca cevapladıkları madde sayısı, madde cevaplamak için harcadıkları ortalama süre ve testin tamamına harcadıkları süre düşük yetenek düzeyindekilere göre daha fazladır.
- Sözel testinde yetenek düzeyi yüksek olanların madde cevaplamak için harcadıkları ortalama süre düşük yetenek düzeyindekilere göre daha fazladır ancak test boyunca cevapladıkları madde sayısı ve teste harcadıkları toplam süre daha azdır.
- Sayısal testinde maddenin zorluk düzeyine göre harcanan süre değişmekte ancak buna bağlı olarak doğru veya yanlış cevaplama durumu değişmemektedir. Öte yandan, herhangi bir maddenin doğru veya yanlış cevaplama durumuna göre harcanan süre değişmektedir. Sözel testinde de benzer bir durum vardır fakat sayısal testinde olduğu kadar bir maddenin doğru veya yanlış cevaplama durumuna göre harcanan süre etkilenmemektedir.
- Cevaplama süresini dikkate almayan MTK ile elde edilen madde parametreleri ile LNCSM ile elde edilen madde parametreleri arasında ilişki olma durumları MTK

modeline göre (Rasch-2PLM) deęişse de süre parametreleri ile madde parametreleri arasında ilişki anlamlıdır.

- LNCSM ile kestirilen hız parametresi ile yetenek deęerleri arasında negatif ilişki çıkmıştır. Maddeye ve testin toplamına harcanan süreler özellikle düşük ve yüksek yetenek grubundaki bireyler için etkilidir. Ayrıca, LNCSM ile MTK modellerinden elde edilen yetenek deęerleri arasında da ilişki elde edilmiştir.

## Öneriler

- Madde parametresi ve yetenek kestirimlerinde LNCSM dışında önerilen dięer madde cevaplama süresini dikkate alan modeller ile de kestirimler yapılabilir.
- ALES'e giren adayların gerçek puanları ile BBT ve LNCSM ile elde edilen yetenek düzeyleri arasındaki ilişkiler incelenebilir.
- MTK'da 3PLM model ile elde edilecek madde parametreleri ve yetenek düzeyleri ile LNCSM sonuçları da karşılaştırılabilir.
- Madde cevaplama süresinin dikkate alındığı BBT uygulaması ile madde cevaplama süresinin dikkate alınmadığı BBT uygulaması aynı bireylere uygulanıp elde edilen yetenek kestirimleri ve maddelere harcanan süreler karşılaştırılabilir.



## Kaynaklar

- ACRE (Accountability and Curriculum Reform Effort). (2010). Computerized Adaptive Testing: How CAT May be Utilized in the Next Generation of Assessments.
- Alkan, V. (2021). *Mesleki Alan İlgi Envanteri'nin bilgisayar ortamında bireye uyarlanmış formunun geliştirilmesi (Yayınlanmamış doktora tezi)*. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Altuner, F. (2019). *Examining the relationship between item statistics and item response time* [Master's Thesis, Mersin University]. Retrieved from <http://tez2.yok.gov.tr/>
- Atar, B. (2011). Tanımlayıcı ve açıklayıcı madde tepki modellerinin TIMSS 2007 Türkiye matematik verisine uyarlanması. *Eğitim ve Bilim*, 36(159).
- Aybek, E.C. (2016). *Kendini Değerlendirme Envanteri'nin Bilgisayar Ortamında Bireye Uyarlanmış Test (BOBUT) Olarak Uygulanabilirliğinin Araştırılması* (Yayınlanmamış doktora tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Aydoğan, İ. (2022). *Okuma Becerilerine Yönelik Yanıt Doğruluğu ve Yanıt Sürelerinin Hiyerarşik Modellenmesi* (Yayınlanmamış doktora tezi). Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Bagus, H. C., Tola, B. ve Tjalla, A. (2022). Achievement tests administration using Computerized Adaptive Testing (CAT) with constrain response time item. *International Journal of Innovation, Creativity and Change*, 16(2).
- Baker, F. B. (1983). *The basic of Item response theory*. Portsmouth, NH: Heinemann.
- Baker, F. B. (2001). *The basics of item response theory*. Tam metin için: <http://ericae.net/irt/baker..>
- Bartram, D. ve Hambleton, R. K. (2006). Computer-based testing and the Internet. *Issues and advantes*. England: John Wiley and Sons.

- Bentler, P. M.ve Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological bulletin*, 88(3), 588.
- Bergstrom, B., Gershon, R. ve Lunz, M. E. (1994). *Computerized adaptive testing: Exploring Examinee response time using Hierarchical liner modeling*. Paper presented at the annul meeting of the National Council on Measurement in Education. New Orleans, LA.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. *Statistical theories of mental test scores*.
- C.A.T. Central (2010), "Operational CAT Testing Programs", <http://www.psych.umn.edu/psylabs/catcentral/operationalcatprograms.htm>, 13.01.2010
- Chang, H. H. ve Ying, Z. (1999). A-stratified multistage computerized adaptive testing. *Applied Psychological Measurement*, 23(3), 211-222.
- Choi, S. W., Grady, M. W.ve Dodd, B. G. (2011). A new stopping rule for computerized adaptive testing. *Educational and Psychological Measurement*, 71(1), 37-53.
- Choi, Y. ve McClenen, C. (2020). Development of adaptive formative assessment system using computerized adaptive testing and dynamic bayesian networks. *Applied Sciences*, 10(22), 8196. doi:10.3390/app10228196
- Čisar, S. M., Radosav, D., Markoski, B., Pinter, R., & Čisar, P. (2010). Computer adaptive testing of student knowledge. *Acta Polytechnica Hungarica*, 7(4), 139-152.
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Psychology press, s. 243.
- Çelik, H. E. ve Yılmaz, V. (2013). Lisrel 9.1 İle Yapısal Eşitlik Modellemesi: Temel Kavramlar Uygulamalar-Programlama (Yenilenmiş 2. Baskı). *Ankara: Anı Yayıncılık*.
- Çıkrıkçı Demirtaşlı, N. (1999). Psikometride yeni ufuklar: bilgisayar ortamında bireye uyarlanmış test. *Türk Psikoloji Bülteni*, 5(13), 31-36.

- Çörtük, M. (2022). *Çok kategorili puanlanan maddelerden oluşan testlerde klasik test kuramı ve madde tepki kuramına dayalı test eşitleme yöntemlerinin karşılaştırılması* (Yayınlanmamış yüksek lisans tezi). Akdeniz Üniversitesi, Eğitim Bilimleri Enstitüsü.
- De Ayala, R. J. (1992). Nominal response model in computerized adaptive testing. *Applied Psychological Measurement*, 16, 327-343.
- De Ayala, R.J. (2009). *The theory and practice of item response theory*. New York: The Guilford Press.
- De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F. ve Partchev, I. (2011). The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software*, 39 (12). doi: 10.18637/jss.v039.i12
- De Beurs, D. P., de Vries, A. L., de Groot, M. H., de Keijser, J. ve Kerkhof, A. J. (2014). Applying computer adaptive testing to optimize online assessment of suicidal behavior: a simulation study. *Journal of Medical Internet Research*, 16(9), e207. doi: 10.2196/jmir.3511
- DeMars, C. (2010). *Item response theory*. New York: Oxford.
- DeMars, C. (2016) *Madde tepki kuramı* (Çev. Ed. Kelecioğlu, H.) Nobel Akademi Yayıncılık, Ankara.
- Demirel, E. (2021). *Ölçme ve değerlendirmede klasik test kuramı ve madde tepki kuramı uygulamaları* (Yayınlanmamış yüksek lisans tezi). Ege Üniversitesi, Fen Bilimleri Enstitüsü.
- Ercikan, K., Julian, M. W., Burket, G. R., Weber, M. M. ve Link, V. (1998). Calibration and scoring of tests with multiple-choice and constructed-response item types. *Journal of Educational Measurement*, 35(2), 137-154.
- Embretson, S. E. ve Reise, S. P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence-Earlbaum Associates.

- Embretson, S. E. ve Reise, S. P. (2013). *Item response theory*. Psychology Press.
- Fox, J. P., Entink, R. K. ve Van der Linden, W. (2007). Modeling of responses and response times with the package cirt. *Journal of Statistical software*, 20, 1-14.
- Fox, J. P., Klotzke, K. ve Simsek, A. S. (2023). R-package LNIRT for joint modeling of response accuracy and times. *PeerJ Computer Science*, 9, e1232.
- Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H. ve Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, 106(3), 608-626. <https://doi.org/10.1037/a0034716>
- Gushta, M. M. (2003). *Standard-setting issues in computerized-adaptive testing*. Paper Prepared for Presentation at the Annual Conference of the Canadian Society for Studies in Education. Halifax. Nova Scotia. May 30th, 2003.
- Gür, R. ve Gülleroğlu, H. D. (2020). Bireye Uyarlanmış Testlerde Madde Kullanım Sıklığı Kontrol Yöntemlerinin Farklı Koşullarda Ölçme Duyarlılığına ve Test Güvenliğine Etkisi. *Eğitim ve Bilim*, 45(202).
- Hambleton, R. K., Swaminathan, H. ve Rogers, H. J. (1991). *Fundamentals of item response theory*. London: Sage Publications
- Hartig, J. ve Höhler, J. (2009). Multidimensional IRT models for the assessment of competencies. *Studies in Educational Evaluation*, 35(2-3), 57-63. <https://doi.org/10.1016/j.stueduc.2009.10.002>
- Hornke, L. F. (2000). Item response times in computerized adaptive testing. *Psicológica*, 21(1), 175-189.
- Istiyono, E., Dwandaru, W. S. B., Setiawan, R. ve Megawati, I. (2020). Developing of computerized adaptive testing to measure physics higher order thinking skills of senior

- high school students and its feasibility of use. *European Journal of Educational Research*, 9(1), 91-101.
- İlgün Dibek, M. (2020). Silent predictors of test disengagement in PIAAC 2012. *Journal of Measurement and Evaluation in Education and Psychology*, 11(4), 430-450. <https://doi.org/10.21031/epod.796626>
- Jöreskog, K. G. ve Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific software international.
- Ju, D. R. (2021). *Explanatory Item Response Time Modeling*. University of Notre Dame.
- Kalender, İ. (2004a). "Bilgisayar Ortamında Bireyselleştirilmiş Testlerin Eğitimde Kullanımı", *XIII. Ulusal Eğitim Bilimleri Kurultayı, İnönü Üniversitesi, Eğitim Fakültesi, Malatya*.
- Kalender, İ. (2011). *Effects of different computerized adaptive testing strategies on recovery of ability*. (Yayınlanmamış doktora tezi). Orta Doğu Teknik Üniversitesi, Fen Bilimleri Enstitüsü.
- Kensinger, E. A., Garoff-Eaton, R. J. ve Schacter, D. L. (2007, May). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language*, 56 (4), 575–591. doi: 10.1016/j.jml.2006.05.004
- Kezer, F. (2013). *Bilgisayar Ortamında Bireye Uyarlanmış Test Stratejilerinin Karşılaştırılması* (Yayınlanmamış doktora tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Kingsbury, G. G. ve Hauser, C. (2004, April). *Computer adaptive testing and the No Child Left Behind Act*. Paper presented at the annual meeting of the American Educational Research Association, San Diego CA. <https://eric.ed.gov/?id=ED491245> adresinden 14.07.2023 tarihinde alınmıştır.
- Lee, Y. H. ve Jia, Y. (2014). Using response time to investigate students' test-taking behaviors in a NAEP computer-based study. *Large-scale Assessments in Education*, 2(1), 1–24. doi: 10.1186/s40536-014-0008-1.

- Michaelides, M. P. ve Ivanova, M. (2022). Response time as an indicator of test-taking effort in PISA: Country and item-type differences. *Psychological Test and Assessment Modeling*, 64(3), 304-338.
- Mills, G. E. ve Gay, L. R. (2019). Educational research: Competencies for analysis and applications. Pearson. One Lake Street, Upper Saddle River, New Jersey 07458.
- Nguyen, T. H., Han, H. R., Kim, M. T. ve Chan, K. S. (2014). An introduction to item response theory for patient-reported outcome measurement. *The Patient-Patient-Centered Outcomes Research*, 7, 23-35.
- Nguyet, P. T. T. ve Daud, M. (2021). Computer Adaptive Test Development To Assess Students' Psychology. *Jurnal Serambi Ilmu*, 22(1), 139-149.
- Ofek Shanny, Y. (2020). *Consequences of Heterogeneous Motivation in Low-Stakes Standardized Assessment Tests* (Doctoral dissertation, University of Haifa (Israel)).
- Özbaşı, D. (2014). *Bilgisayar okuryazarlığı testinin bilgisayar ortamında bireye uyarlanmış test olarak uygulanabilirliğine ilişkin bir araştırma* (Yayınlanmamış doktora tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Paulhus, D. L. (2002). *Socially desirable responding: The evolution of a construct*. In H. I. Braun, D. N. Jackson ve D. E. Wiley (Ed.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.
- Raiche, G.ve Blais, J. G. (2002). Practical Considerations about Expected A Posteriori Estimation in Adaptive Testing: Adaptive A Priori, Adaptive Correction for Bias, and Adaptive Integration Interval.
- Qian, H., Staniewska, D., Reckase, M. ve Woo, A. (2016). Using response time to detect item preknowledge in computer-based licensure examinations. *Educational Measurement: Issues and Practice*, 35(1), 38-47.

- Schnipke, D.L. ve Scrams, D.J. (1997). Modeling item response times with a two-state mixture model: A new method of measuring speededness. *Journal of Educational Measurement*, 34, 213-232.
- Selçuk, E. (2023). *Klasik ve Bayeşçi yaklaşımlara göre madde tepki kuramı parametre kestirimlerinin farklı simülasyon koşullarında karşılaştırılması* (Yayınlanmamış doktora tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Smith, R. W. (2000). *An exploratory analysis of item parameters and characteristics that influence item level response time* (Doctoral dissertation, University of Nebraska-Lincoln, 2000). Dissertation Abstract International, 61, 05A.
- Song, T. (2010). *The effect of fitting a unidimensional IRT model to multidimensional data in content-balanced computerized adaptive testing*. Michigan State University.
- Steffen, M. ve Way, D. W. (1999, April). Testing-taking strategies in computerized adaptive testing. Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, Canada.
- Suh, H. (2010). *A study of Bayesian estimation and comparison of response time models in item response theory* (Doctoral dissertation, University of Kansas).
- Susmann, H. (2018). Item Response Theory (IRT) Models. Retrieved July 18, 2023, from <https://observablehq.com/@herbps10/item-response-theory-irt-models>
- Şahin, M. D. (2017). *Gerçek ve üretilmiş veri setlerinde çok boyutlu bireyselleştirilmiş bilgisayarlı test uygulama sonuçlarının incelenmesi* (Yayınlanmamış doktora tezi). Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Şahin Kürşad, M. (2023). The effects of different item selection methods on test information and test efficiency in computer adaptive testing. *Journal of Measurement and Evaluation in Education and Psychology*, 14(1), 33-46. <https://doi.org/10.21031/epod.1140757>

- Şenel, S. (2017). *Bilgisayar ortamında bireye uyarlanmış testlerin görme engelli öğrencilere uygunluğunun incelenmesi* (Yayınlanmamış doktora tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Tancoš, M., Chvojka, E., Jabůrek, M. ve Portešová, Š. (2023). Faster≠ Smarter: Children with higher levels of ability take longer to give incorrect answers, Especially when the task matches their ability. *Journal of Intelligence*, 11(4), 63. <https://doi.org/10.3390/jintelligence11040063>
- Tekin, H. (2000). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınları.
- Tian, J. Q., Miao, D. M., Zhu, X.ve Gong, J. J. (2007). An Introduction to the Computerized Adaptive Testing. *Online Submission*, 4(1), 72-81.
- Thissen, D. (1983). Timed testing: An approach using item response theory. In D. J. Weiss (Ed.), *New horizons in testing: Latent trait test theory and computerized adaptive testing*. New York: Academic Press.
- Thompson, N. A. (2010). Adaptive testing: Is it right for me. *Minesota: University of Cincinnati. Assessment Systems Corporation*. [https://assess.com/docs/Thompson \(2010\) - Adaptive Testing Right.pdf](https://assess.com/docs/Thompson (2010) - Adaptive Testing Right.pdf) adresinden, 10 Temmuz 2023'te alınmıştır.
- Türkoğuz, S. (2020). Comparison of yhreshold balues of Three-Tier Diagnostic and Multiple-Choice Tests based on response time. *Anatolian Journal of Education*, 5(2), 19-36. <https://doi.org/10.29333/aje.2020.522>.
- Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F.ve Mislevy, R. J. (2000). *Computerized adaptive testing: A primer*. Routledge.
- Wainer, H., Bradlow, E. T. ve Wang, X. (2007). *Testlet response theory and its applications*. Cambridge University Press.
- Wang, M. (2017). *Characteristics of item response time for standardized achievement assessments* (Doctoral dissertation, University of Iowa).

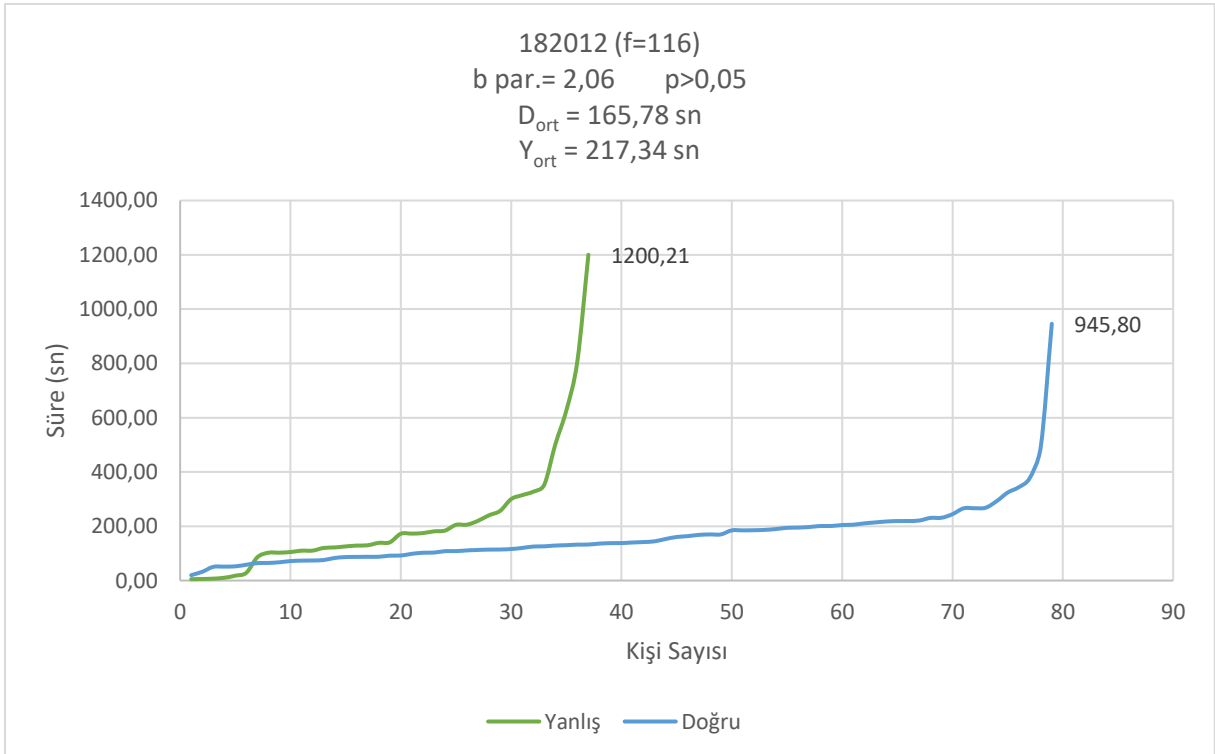
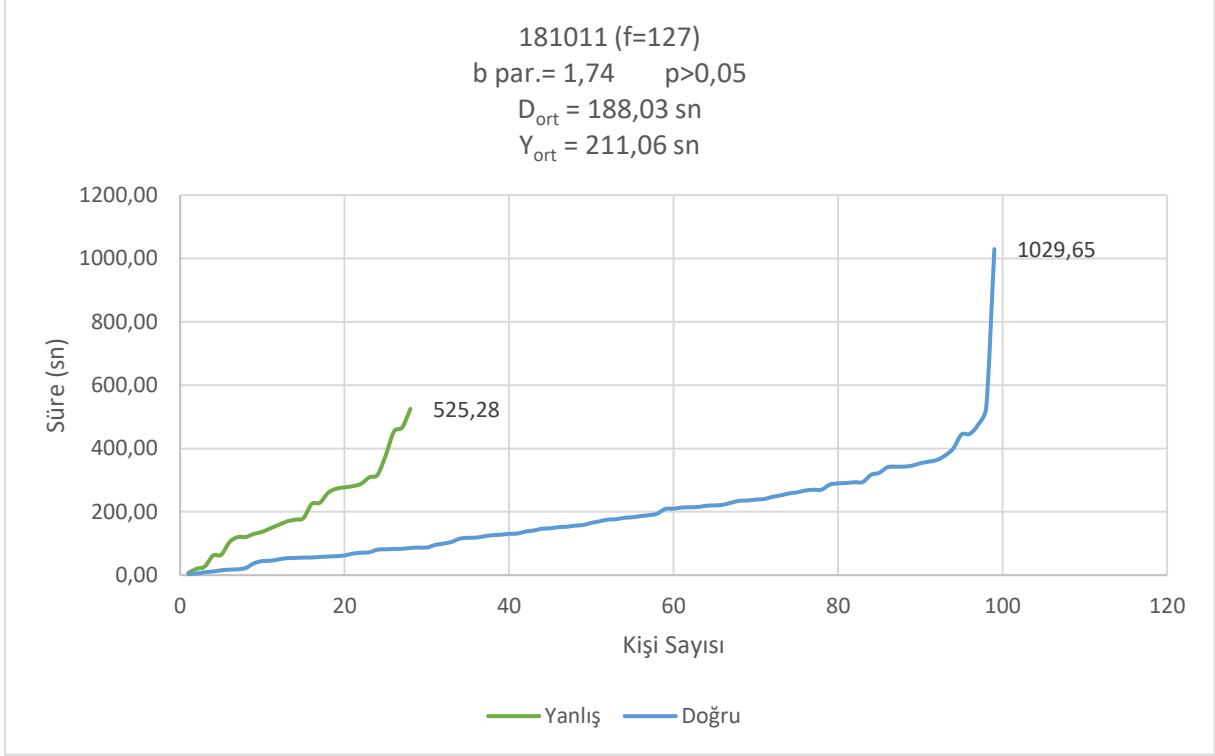


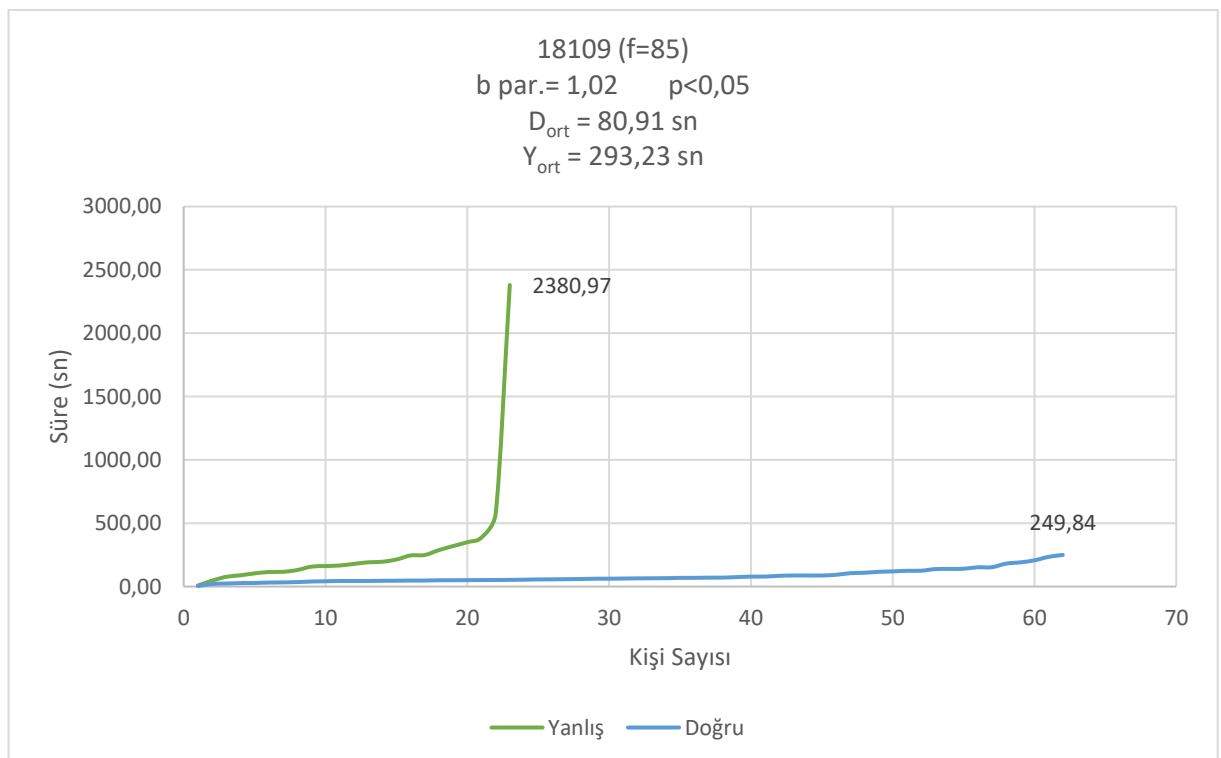
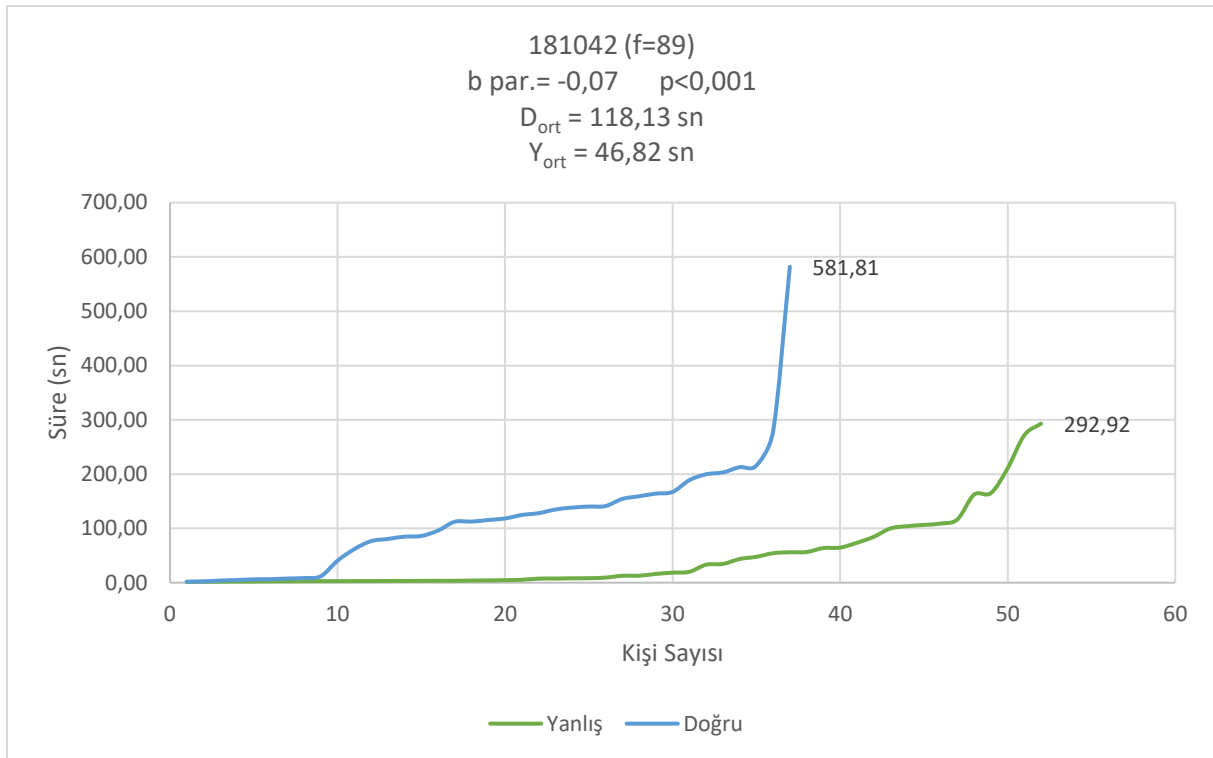
- Wang, T. ve Hanson, B. A (2005). Development and calibration of an item response model that incorporates response time. *Applied Psychological Measurement*, 29, 323-339.
- Weiss, D. J. (2004). Computerized adaptive testing for effective and efficient measurement in counseling and education. *Measurement and Evaluation in Counseling and Development*, 37(2), 70-84.
- Weiss, D. J. (2011). Better data from better measurements using computerized adaptive testing. *Journal of Methods and Measurement in the Social Sciences*, 2(1), 1-27.
- Wise, S. L. ve DeMars, C. E. (2006). An application of item response time: The effort-moderated IRT model. *Journal of Educational Measurement*, 43(1), 19-38.
- Wise S. L. ve DeMars C. E. (2010). Examinee noneffort and the validity of program assessment results. *Educational Assessment*, 15(1), 27-41.
- Wise, S. L. ve Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, 18(2), 163-183.  
[https://doi.org/10.1207/s15324818ame1802\\_2](https://doi.org/10.1207/s15324818ame1802_2)
- van der Linden, W. J. (2006). A lognormal model for response times on test items. *Journal of Educational and Behavioral Statistics*, 31(2), 181-204.
- Van der Linden, W. J. (2007). A hierarchical framework for modeling speed and accuracy on test items. *Psychometrika*, 73, 287-308.
- Van der Linden, W. J. (2008). Using response times for item selection in adaptive testing. *Journal of Educational and Behavioral Statistics*, 33(1), 5-20.
- Van der Linden, W. J. ve Glas, C. A. (Ed.). (2000). *Computerized adaptive testing: Theory and practice*. Springer Science ve Business Media.
- Van der Linden, W. J. ve Guo, F. (2008). Bayesian procedures for identifying aberrant response-time patterns in adaptive testing. *Psychometrika*, 73(3), 365-384.

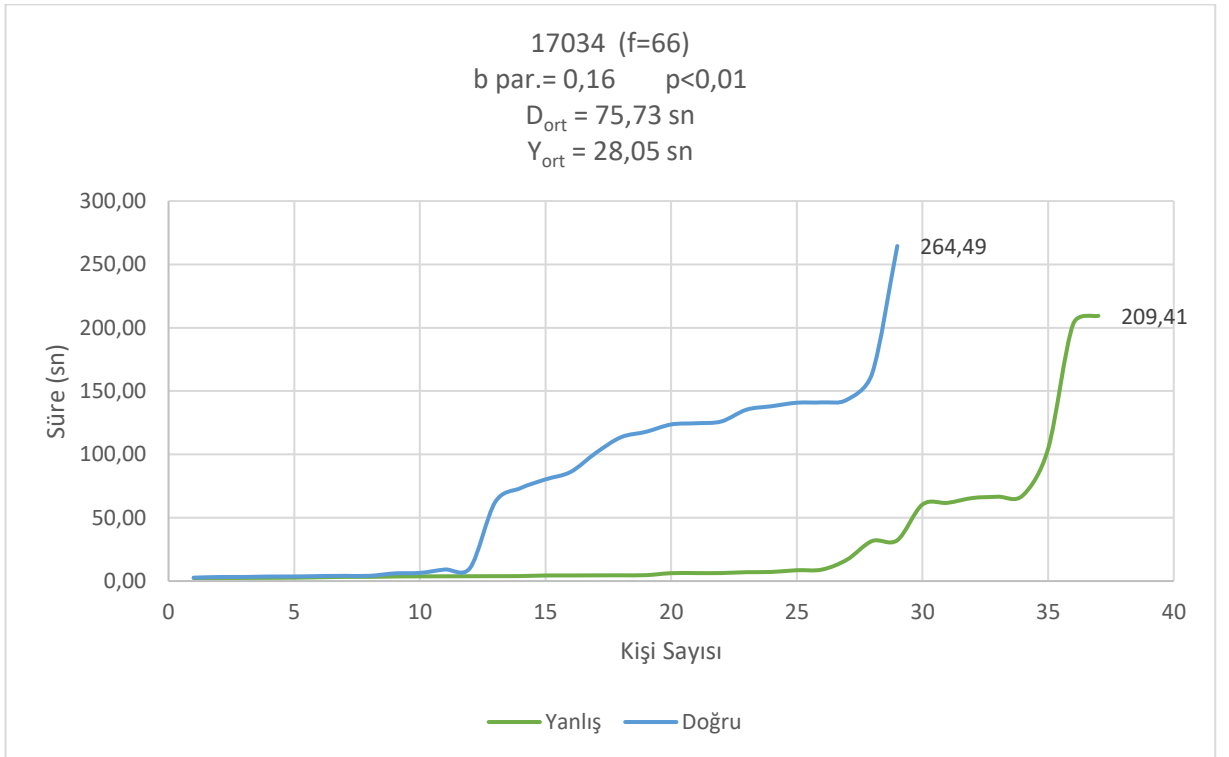
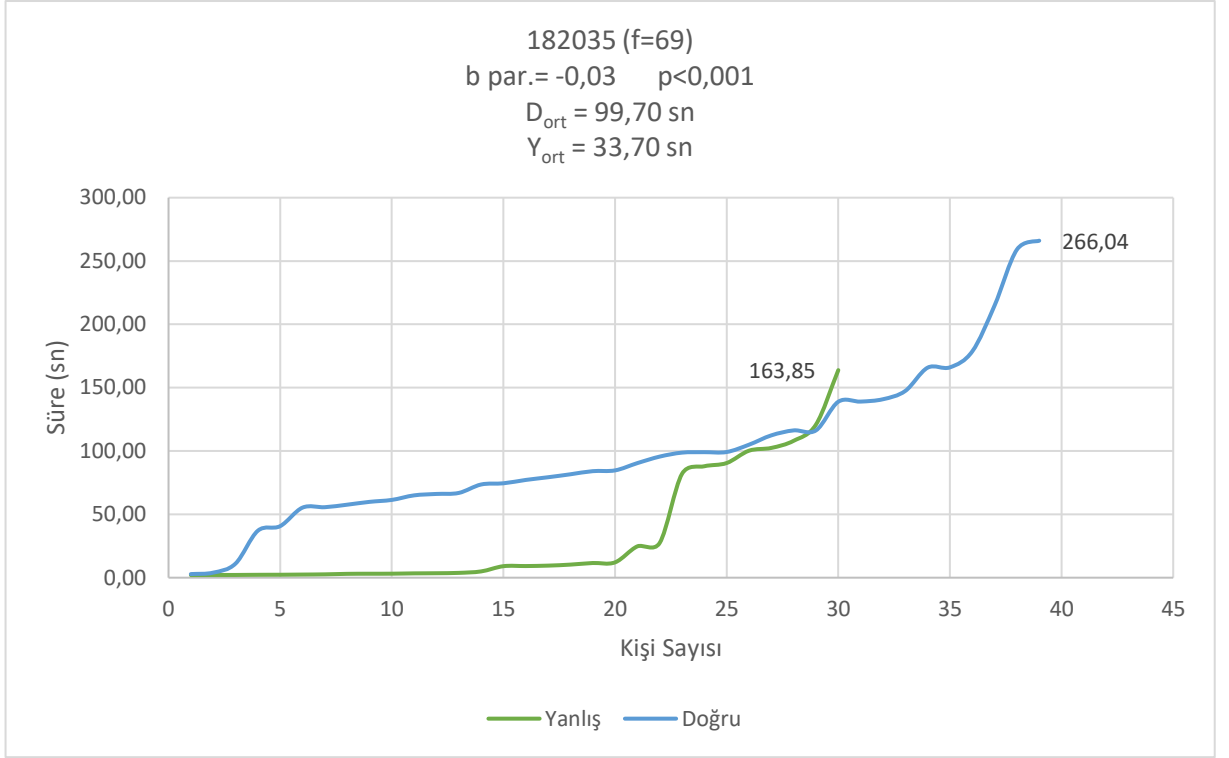
- Yalçın, S. (2022). Examining students' item response times in eTIMSS according to their proficiency levels, self-confidence, and item characteristics. *Journal of Measurement and Evaluation in Education and Psychology*, 13(1), 23-39. <https://doi.org/10.21031/epod.999545>
- Yaşar, M. ve Aybek, E. C. (2019). Üniversite Öğrencileri için Bir Yılmazlık Ölçeğinin Geliştirilmesi: Madde Tepki Kuramı Temelinde Geçerlilik ve Güvenilirlik Çalışması. *İlköğretim Online*, 18(4). doi:10.17051/ilkonline.2019.635031
- Yavuz, H. Ç. (2019). The effects of log data on students' performance. *Journal of Measurement and Evaluation in Education and Psychology*, 10(4), 378-390. <https://doi.org/10.21031/epod.564232>
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125-145. <https://doi.org/10.1177/014662168400800>
- Yıldırım, H. H., Çömlekoğlu, G. ve Berberoğlu, G. (2003). Milli Eğitim Bakanlığı özel okullar sınavı verilerinin madde tepki kuramı modellerine uyumu. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 24(24).

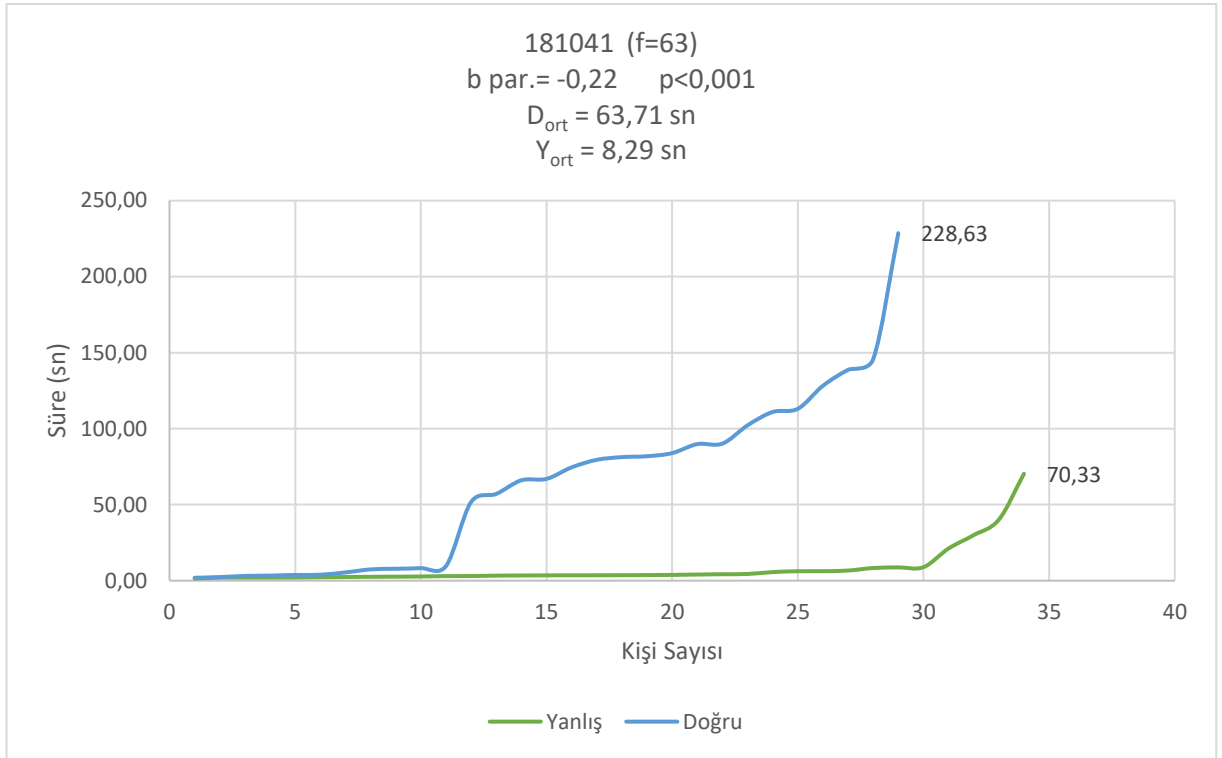
## EKLER

**Ek-A: Sayısal Testinde Madde Cevaplama Süresinin Maddenin Doğru veya Yanlış Cevaplanma Durumuna Göre Değişim Grafikleri**

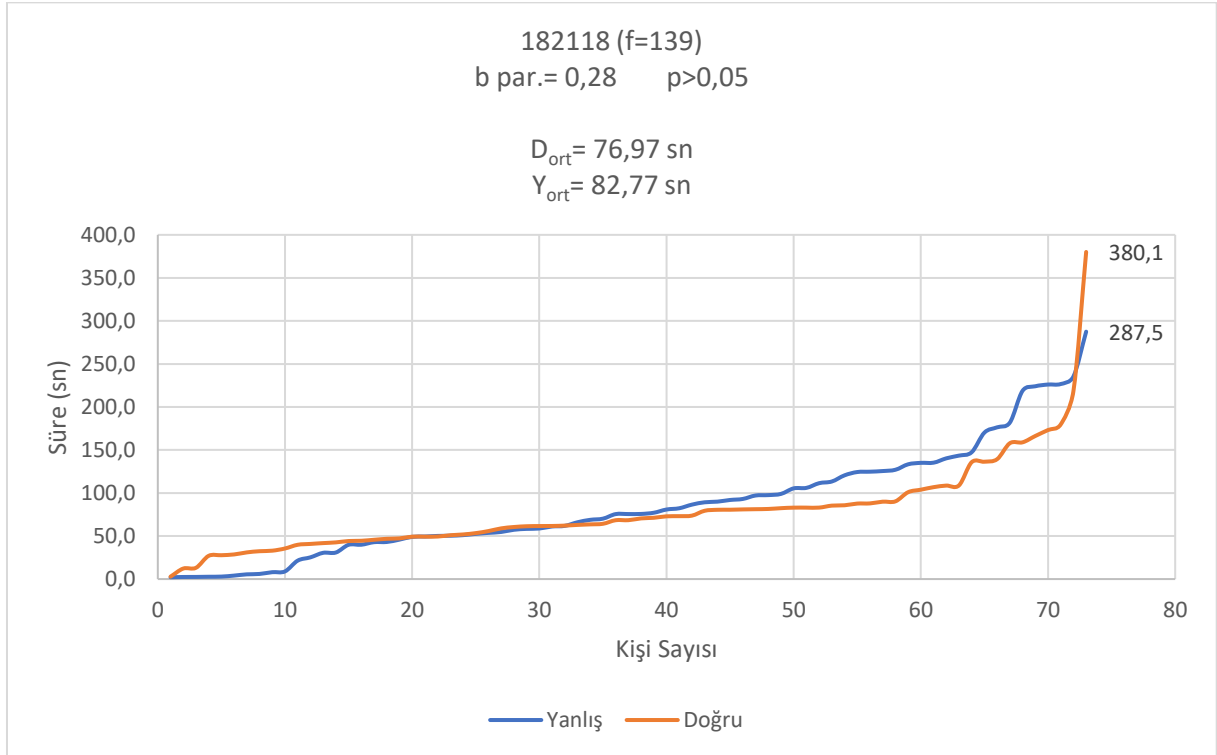
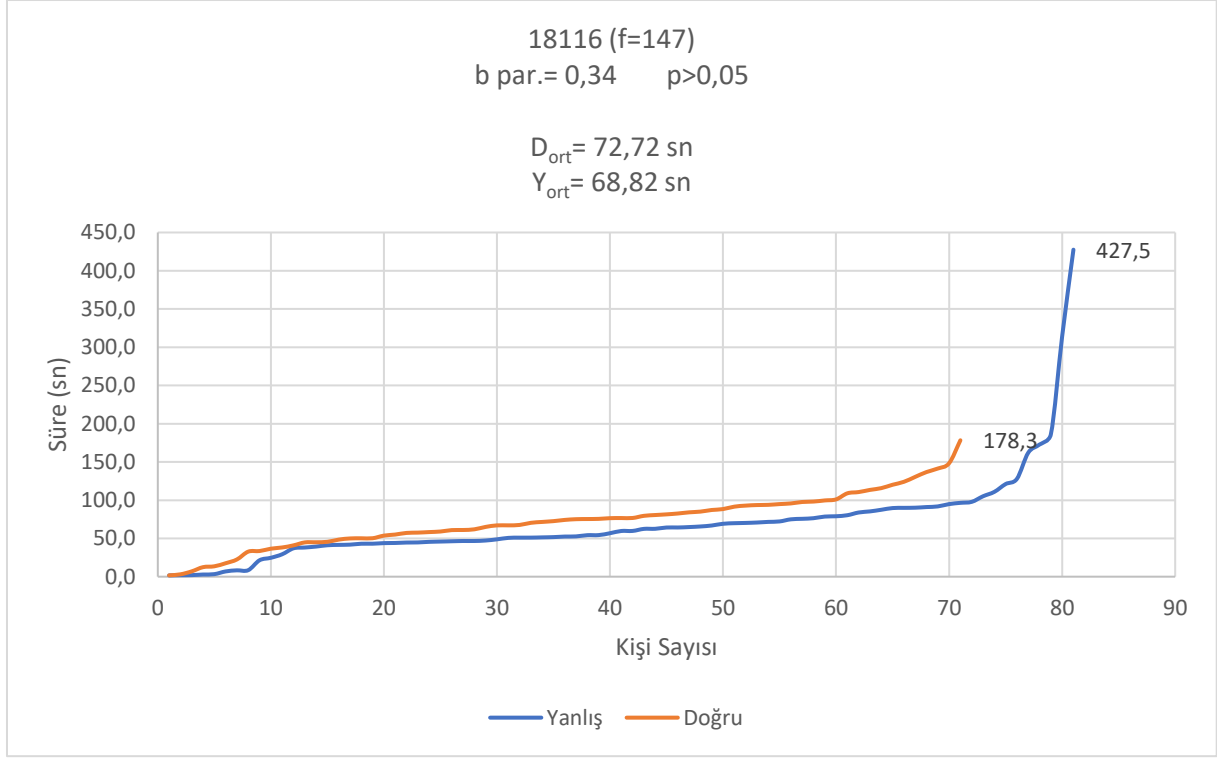


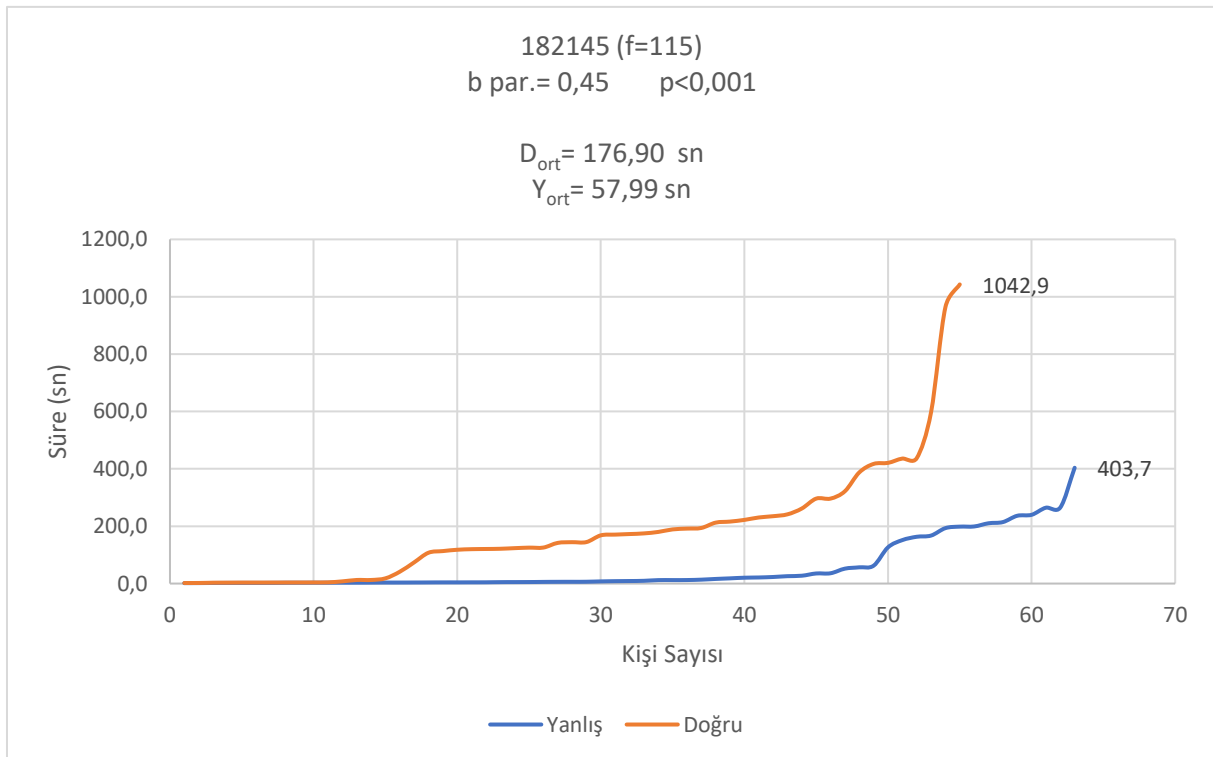
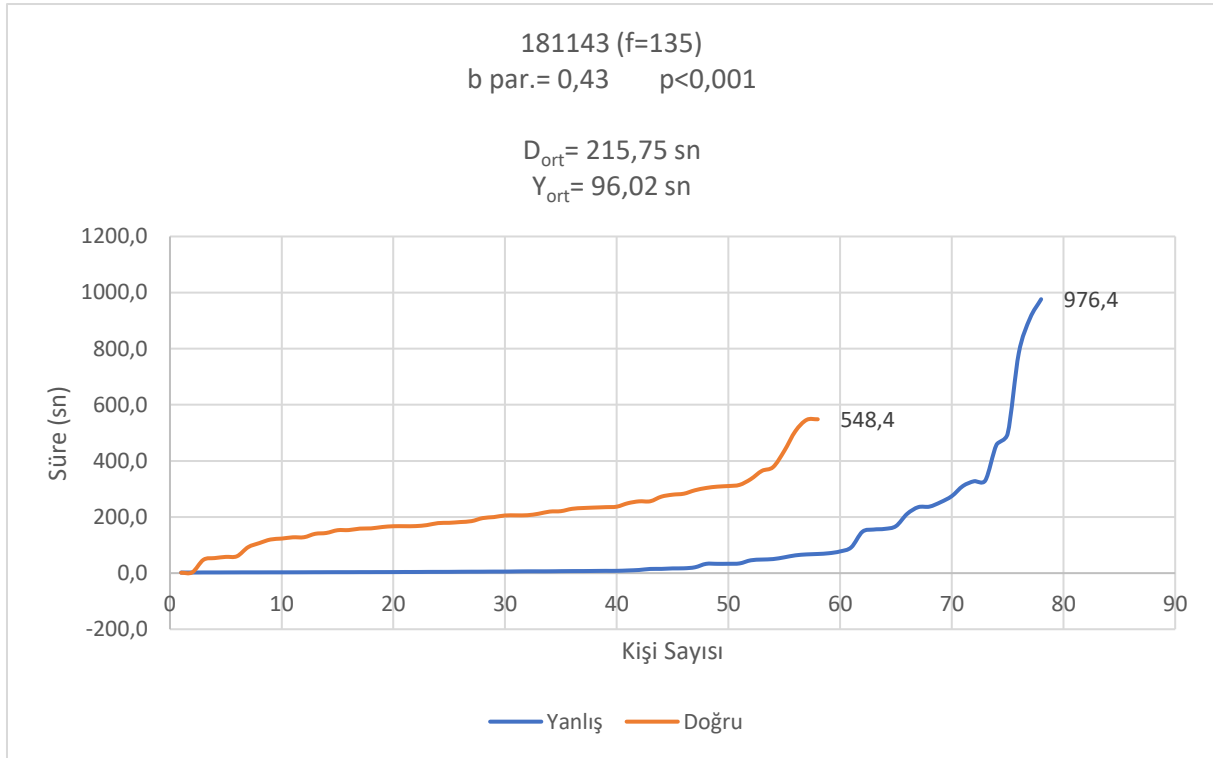




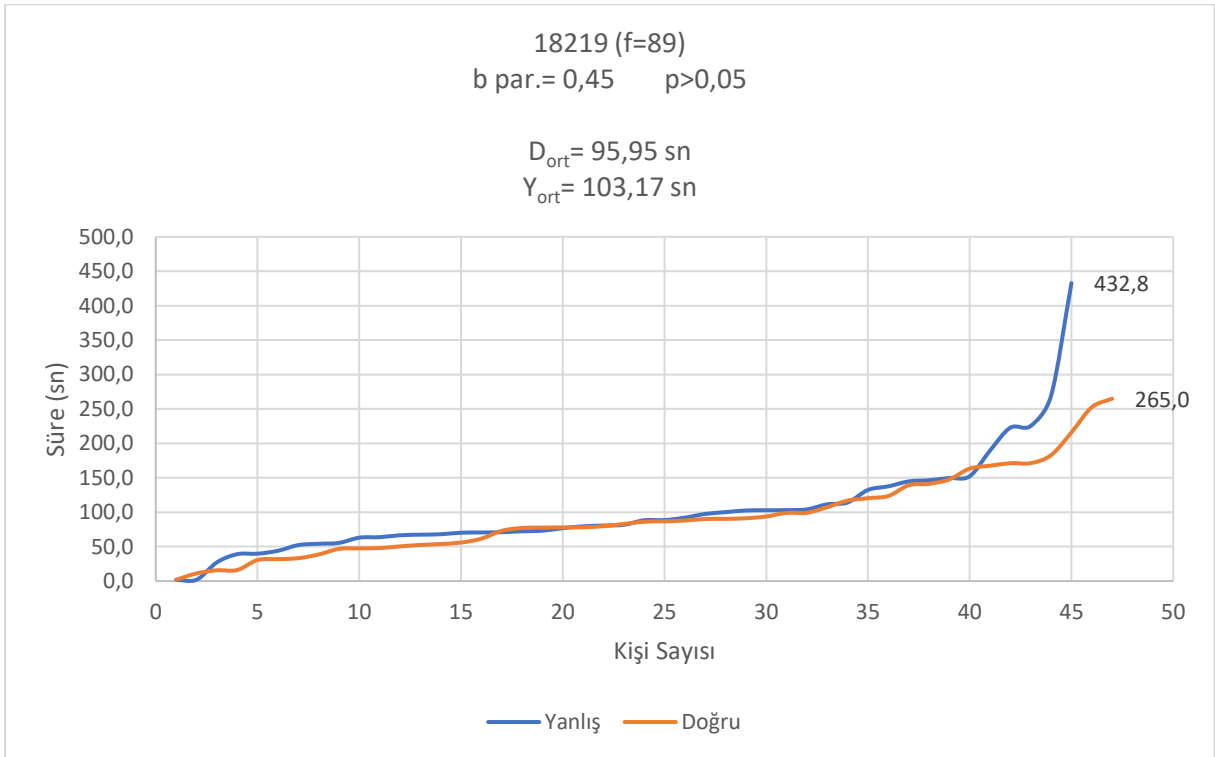
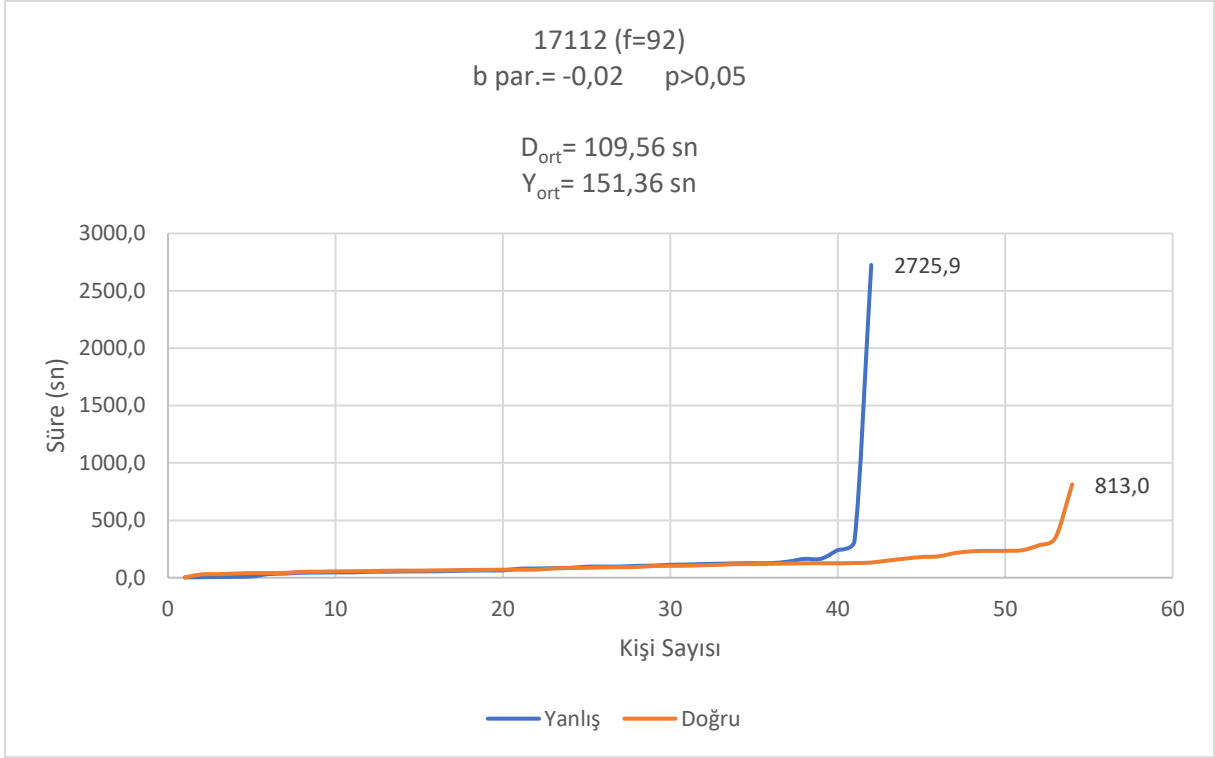


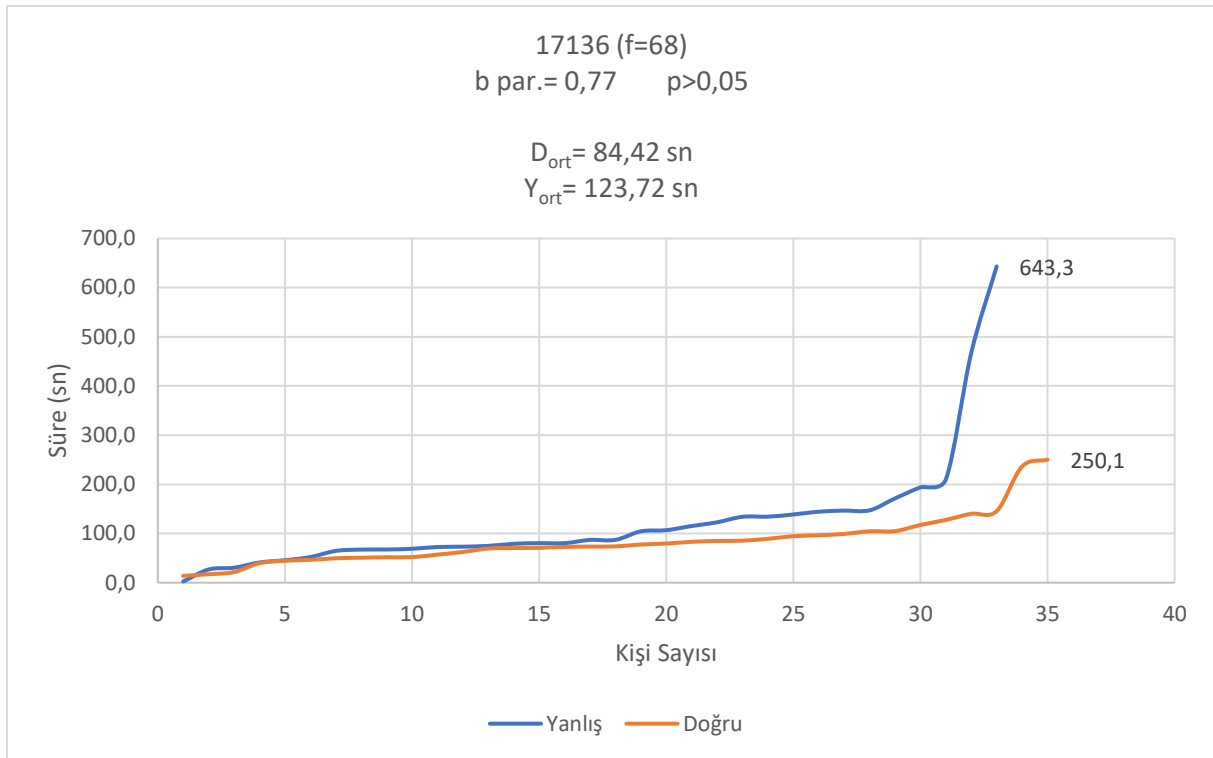
**Ek-B: Sözel Testinde Madde Cevaplama Süresinin Maddenin Doğru veya Yanlış Cevaplanma Durumuna Göre Değişim Grafikleri**











### Ek-C: LNCSM Sonucu Elde Edilen Madde Parametrelerinin Betimsel Analiz Sonuçları

#### Sayısal Testinde LNCSM Sonucu Elde Edilen Madde Parametreleri Betimsel Analiz

##### Sonuçları

	En	En	$\bar{X}$	ss	Çarpıklık		Basıklık	
	düşük değer	yüksek değer			İstatistik	SH	İstatistik	SH
M. Ayırt Ediciliği	0,23	1,84	1,18	0,35	-0,35	0,18	-0,55	0,35
Madde Güçlüğü	-0,47	0,30	-0,04	0,14	-0,46	0,18	0,25	0,35
S. Ayırt Ediciliği	0,70	1,16	1,01	0,06	-1,85	0,18	6,67	0,35
Süre Yoğunluğu	3,44	4,74	4,24	0,18	-0,56	0,18	2,11	0,35

Not:  $n = 192$ . M: Madde; S: Süre

#### Sözel Testinde LGIRT Model Sonucu Elde Edilen Madde Parametreleri Betimsel Analiz

##### Sonuçları

	En	En	$\bar{X}$	ss	Çarpıklık		Basıklık	
	düşük değer	yüksek değer			İstatistik	SH	İstatistik	SH
M. Ayırt Ediciliği	0,19	2,03	1,10	0,20	-0,05	0,18	4,06	0,35
Madde Güçlüğü	-0,47	1,18	-0,01	0,16	2,00	0,18	14,84	0,35
S. Ayırt Ediciliği	0,59	1,23	1,02	0,09	-2,19	0,18	7,60	0,35
Süre Yoğunluğu	3,46	4,37	3,81	0,11	1,29	0,18	5,20	0,35

Not:  $n = 192$ . M: Madde; S: Süre

## Ek-D: LNCSM Sonucu Elde Edilen Yetenek ve Hız Parametrelerinin Betimsel Analiz

### Sonuçları

#### Sayısal Testinde LNCSM Model Sonucu Elde Edilen Yetenek ve Hız Değerlerinin Betimsel

##### Analiz Sonuçları

	En	En	$\bar{X}$	ss	Çarpıklık		Basıklık	
	düşük değer	yüksek değer			İstatistik	SH	İstatistik	SH
LNCSM_yetenek	-1,62	1,03	-0,01	0,61	-0,75	0,16	-0,22	0,31
LNCSM_hız	-1,92	3,14	-0,01	1,12	1,28	0,16	0,71	0,31

Not:  $n = 238$ .

#### Sözel Testinde LNCSM Model Sonucu Elde Edilen Yetenek ve Hız Değerlerinin Betimsel

##### Analiz Sonuçları

	En	En	$\bar{X}$	ss	Çarpıklık		Basıklık	
	düşük değer	yüksek değer			İstatistik	SH	İstatistik	SH
LNCSM_yetenek	-1,20	0,86	0,00	0,43	-0,75	0,17	-0,29	0,34
LNCSM_hız	-1,55	3,09	0,00	1,12	1,24	0,17	0,35	0,34

Not:  $n = 203$ .

## Ek-E: Arařtırma Etik Komisyonu Onay Bildirimi



T.C.  
HACETTEPE ÜNİVERSİTESİ REKTÖRLÜĞÜ  
Rektörlük



Sayı : E-35853172-300-00002680102  
Konu : Emine GÜLEN ULUSOY Hk. (Etik Komisyon İzni)

8.02.2023

### EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 02.01.2023 tarihli ve E-51944218-300-00002605043 sayılı yazımız.

Enstitünüz Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme doktora programı öğrencilerinden **Emine Gülen ULUSOY**'un **Prof. Dr. Selahattin GELBAL** danışmanlığında yürüttüğü "**Bireyselleştirilmiş Bilgisayar Uyarlamalı Testlerde Madde Cevaplama Süresinin Yetenek Düzeyi ve Madde Parametrelerine Etkisi**" başlıklı tez çalışması, Üniversitemiz Senatosu Etik Komisyonunun **10 Ocak 2023** tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini rica ederim.

Prof. Dr. Serhat ÜNAL  
Rektör Yardımcısı

Bu belge güvenli elektronik imza ile imzalanmıştır.

Belge Doğrulama Kodu: 8F1B1995-DFCB-472A-8083-0BF805EB6879

Belge Doğrulama Adresi: <https://www.turkiye.gov.tr/hu-eby5>

Adres: Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara  
E-posta: yazind@hacettepe.edu.tr İnternet Adresi: www.hacettepe.edu.tr Elektronik  
Ağ: www.hacettepe.edu.tr  
Telefon: 0 (312) 305 3001-3002 Faks: 0 (312) 311 9992  
Kep: hacettepeuniversitesi@hs01.kep.tr

Bilgi için: Duygu Didem İLERİ  
Bilgisayar İşletmeni  
Telefon: .



**EK-F: Etik Beyanı**

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- \* tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- \* görsel, işitsel ve yazılı bütün bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- \* başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- \* atıfta bulunduğum eserlerin bütününe kaynak olarak gösterdiğimi,
- \* kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- \* bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

...../...../.....

(İmza)

Emine Gülen ULUSOY

**EK-G: Doktora Tez Çalışması Orijinallik Raporu**

28/09/2023

HACETTEPE ÜNİVERSİTESİ  
Eğitim Bilimleri Enstitüsü  
Eğitim Bilimleri Ana Bilim Dalı Başkanlığına,

Tez Başlığı: Bilgisayarda Bireyselleştirilmiş Testlerde Madde Cevaplama Süresi Yetenek Düzeyi ve Madde Parametreleri Arasındaki İlişkinin İncelenmesi

Yukarıda başlığı verilen tez çalışmamın tamamı (kapak sayfası, özetler, ana bölümler, kaynakça) aşağıdaki filtreler kullanılarak **Turnitin** adlı intihal programı aracılığı ile kontrol edilmiştir. Kontrol sonucunda aşağıdaki veriler elde edilmiştir:

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Oranı	Gönderim Numarası
28/09/2023	98	123004	31/07/2023	%13	2179432041

Uygulanan filtreler:

1. Kaynaklar hariç
2. Alıntılar dâhil
3. 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esaslarını inceledim ve çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan eder, gereğini saygılarımla arz ederim.

**Ad Soyadı:** Emine Gülen ULUSOY

**Öğrenci No.:** N11147061

**Ana Bilim Dalı:** Eğitim Bilimleri

İmza

**Programı:** Eğitimde Ölçme ve Değerlendirme

**Statüsü:**  Y.Lisans  Doktora  Bütünleşik Dr.

**DANIŞMAN ONAYI**

UYGUNDUR.

(Prof. Dr., Selahattin GELBAL, İmza)

## EK-H: Dissertation Originality Report

28/09/2023

HACETTEPE UNIVERSITY  
Graduate School of Educational Sciences  
To The Department of Educational Sciences

Thesis Title: Investigating the Correlations Between Response Time, Ability Level and Item Parameters in Computer Adaptive Test

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
28/09/2023	98	123004	31/07/2023	%13	2179432041

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

**Name Lastname:** Emine Gülen ULUSOY

**Student No.:** N11147061

**Department:** Educational Sciences

**Program:** Educational Measurement and Evaluation

**Status:**  Masters  Ph.D.  Integrated Ph.D.

Signature

### ADVISOR APPROVAL

APPROVED  
(Prof. Dr., Selahattin GELBAL, Signature)



## EK-I: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**" kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü/Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir. <sup>(1)</sup>
- Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren ... ay ertelenmiştir. <sup>(2)</sup>
- Tezimle ilgili gizlilik kararı verilmiştir. <sup>(3)</sup>

..... / ..... / .....

(imza)

Emine Gülen ULUSOY

---

"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"

- (1) Madde 6.1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezinerişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6.2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internette paylaşılması durumunda 3 şahıslara veya kurumlara haksız kazanç; imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanın önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7.1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir\*. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.  
Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir  
\*Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir

